CCA External Projects 2016-17

1. 'Use of change-point tools and techniques to identify when a significant change has occurred in noisy signals, in order to identify incipient equipment problems BEFORE failure occurs', Artesis LLP
2. 'Learning the Impact of Freight on Global Commodity Markets', Cantab
3. 'Recommended engine implementation for the development of tailored and revenue-driven promotions', Costa
4. 'Finding Structure, Symmetries and Dynamics in Graphs', DSTL
5. 'Attenuation correction in SPECT CT images', GSK
6. 'Model-based experimental design in imaging studies on animals: how much can we reduce animal numbers?', GSK
7. 'Multi-parametric Textural Analysis of Cartilage', GSK
8. 'A source estimation inverse problem in oilfield drilling: Correcting surface arrival times of drilled cuttings for the effects of dispersion during transport up the well', Schlumberger
9. 'Discrete Fourier transform over regular triangle or tetrahedron', Schlumberger
10. 'Fluid-Fluid Displacement in Boreholes', Schlumberger
11. 'Fully Bayesian analysis of multi-exponential decays in low resolution NMR', Schlumberger

First Year PhD project - Artesis LLP Project Outline

(External Project partners should submit a one to two page description of the project which includes the background, some key literature references, problem, scope, project aims and expectations.)

Project Title:
Use of change-point tools and techniques to identify when a significant change has occurred in noisy signals, in order to identify incipient equipment problems BEFORE failure occurs.

PMP student
TBA

Artesis Contact Name
Geoff Walker

Contact Email
Geoff.walker@artesis.co.uk

Organisation Name
Artesis LLP

Organisation Address
St John's Innovation Centre, Cowley Road, Cambridge, CB4 0WS

Period of the Project
November 2016 – March 2017
8 weeks elapsed, 4 man weeks PhD student input.

Background
Artesis LLP are re-defining equipment healthcare. Artesis LLP helps a wide range of industry to operate more efficiently, saving wasted expense and avoiding disruption from unexpected downtime by identifying and diagnosing incipient problems in industrial equipment, allowing intervention before problems become critical. This is done by analysing the distortions on the waveform of the 3-phase AC current drawn by the motor that is driving the equipment. This current is affected by a wide range of equipment phenomena – including mechanical and operational factors as well as electrical factors – so with the right analytical techniques it can be used to provide a rather precise diagnostic insight into the health of the equipment. We use a whole series of mathematical techniques to do this, some of which are standard techniques like Signal Processing techniques, Fast Fourier Transforms etc, and others that are rather more bespoke, involving System Identification algorithms. This summer a PMP project student, Will Boulton, carried out a review of our end-to-end mathematical processing, identifying areas of opportunity for improvement in the techniques used, and developing some specific solutions.

The area of focus for this First year PhD project will be on one (or possibly two) of those areas identified, specifically including the use of change point tools and techniques to identify when a significant change has occurred, and then to identify the nature of the change, leading to being able to provide the most appropriate advice about what should be done and when to do it. This is not a trivial task because the signals can be noisy, and can be affected by a number of factors that are not indicators of equipment problems. It is expected that use will

be made of the techniques published by John Aston on the website http://www.changepoint.info

Key Literature References
- A good background to the specific environment is the phase 1 report by last summer's PMP student: Will Boulton's report is available on request from geoff.walker@artesis.co.uk

- Lots of useful information can be found on the change point website – www.changepoint.org; specific references are listed on this site at http://www.changepoint.info/publications.html

Of less direct relevance to this study, but of great relevance to the wider approaches used to generate the spectra in question can be found in:

- Principles of System Identification: Theory and Practice Hardcover – 22 Dec 2014 by Arun K. Tangirala

And wider general background to the overall application area (not required for this study, apart from setting the context):
- Introduction to Machinery Analysis and Monitoring - John Mitchell
- An Introduction to Predictive Maintenance - John Mobley

Problem and Scope
The problem is:

to identify when a genuinely significant change has occurred in a noisy signal.

The main signals of interest in this project are spectra of the non-linear components of the current drawn by the electric motor driving an item of equipment. The sort of equipment might typically be a pump, fan or compressor. This sort of equipment concerned might for example be a critical pump providing water supply to a district, or an air conditioning fan supplying clean air to an operating theatre, or a chiller compressor supplying cooling to a critical data centre; in the future it could well be a motor driving an electric car. In all of these situations, allowing the equipment to fail could have highly undesirable consequences.

For each item of equipment being monitored, a new spectrum is generated every minute or so. The shape and level of the spectrum indicate phenomena related to the behaviour of the system, so changes in the shape and level of the spectrum indicate changes occurring in the equipment, which is typically an indication of a developing problem. We want to be able to identify this developing problem before it leads to complete and catastrophic failure of the equipment, and to be able to provide advice on the nature of the problem and how best to deal with it.

There is a well understood correlation between the frequency at which peaks occur on the spectrum and the phenomenon causing it – so for example, a phenomenon such as unbalance or misalignment will typically create a signal at 1 x rotational speed of the equipment. Other phenomena, such as deterioration of rolling element bearings, have their own specific characteristic frequencies. So once a peak is seen to be growing, it is relatively straightforward to generate a diagnosis of the cause, and to provide advice on the appropriate response or corrective action.

However, the signals that generate the spectrum are subject to noise, and so it is not a trivial matter to identify when a particular peak has genuinely started to change its level. Given the noisy nature of the data it is equally important to avoid false alarms as to avoid missed alarms.

And whilst some phenomena show up simply as a single peak, others can show up as a family of harmonics of peaks, and others still can show up as a more general change in shape of the spectrum – such as a broad "hump" in the spectrum, rather than a single sharp peak. The identification of a change needs to be able to recognise all of these different types of behaviour.

Once a genuine change has been identified, there is then a need to be able to generate a trend plot of the level of this phenomenon, to give an indication of the rate of deterioration and hence provide some idea of how rapidly action needs to be taken to correct the underlying problem.

Project Aims and Expectations

The overall aim is to develop and demonstrate a robust method for identifying significant changes in real-world data, that combines a high success rate in raising genuine alarms with a high success rate of avoiding false alarms.

Key aspects expected in this project:

- Real world data will be provided by Artesis. Some of this data will have known developing fault phenomena which can be used to test out the tools and techniques; other data to be supplied will be as yet unassessed, in which we wish to find developing phenomena;
- Background support to the work done in the previous PMP project will be available from the PMP student, Will Boulton;
- Use should be made of the tools and techniques available on changepoint.info website and elsewhere, and support should be sought from staff involved in this area;
- The final output should be to demonstrate successful identification of change on real data with a working set of algorithms encoded in appropriate software (eg Matlab / C++ / other). This output should include an assessment of the likely level of false positives and false negatives in the conclusions.

Some of the steps involved are expected to be:

1. Review the range of tools and techniques available on changepoint.info and elsewhere;
2. Think about the objective and plan out an approach using appropriate tools;
3. Use an appropriate subset of change point tools on real-world data to see what outputs can be generated;
4. Evaluate the potential identification rate of false positives and false negatives based on the nature of the tools and the data, and compare this theoretical "best outcome" rate with what has been generated in step 3;
5. Iterate the selection or configuration of the tools and techniques to obtain the best possible balance of accuracy, effort and speed;
6. Apply the chosen technique to new data to see what previously un-identified phenomena can be found;

7. Write up the findings and conclusions with recommendations of any additional work that would still be required to turn this into a usable day-to-day method.

LEARNING THE IMPACT OF FREIGHT ON GLOBAL COMMODITY MARKETS

Project Supervisor: Dr Chris Longworth, Cantab Capital Partners LLP

CCIMI: Dr Carola-Bibiane Schönlieb

BACKGROUND

Recent research at Cantab has focused on using information contained in novel datasets, such as power consumption and weather forecast data, to predict movements in commodity prices. Unlike traditional financial inputs, such as price or yield curve changes, the nature of the relationship between this data and resulting price move is often idiosyncratic and can vary significantly across markets. While it may be possible to specify the expected relationship a-priori, this typically requires specialist knowledge of individual markets and may not be robust to changes in market structure. A promising alternative has been to apply machine learning techniques such as dynamic bayesian networks[1], and neural networks[2] to dynamically identify tradable relationships.
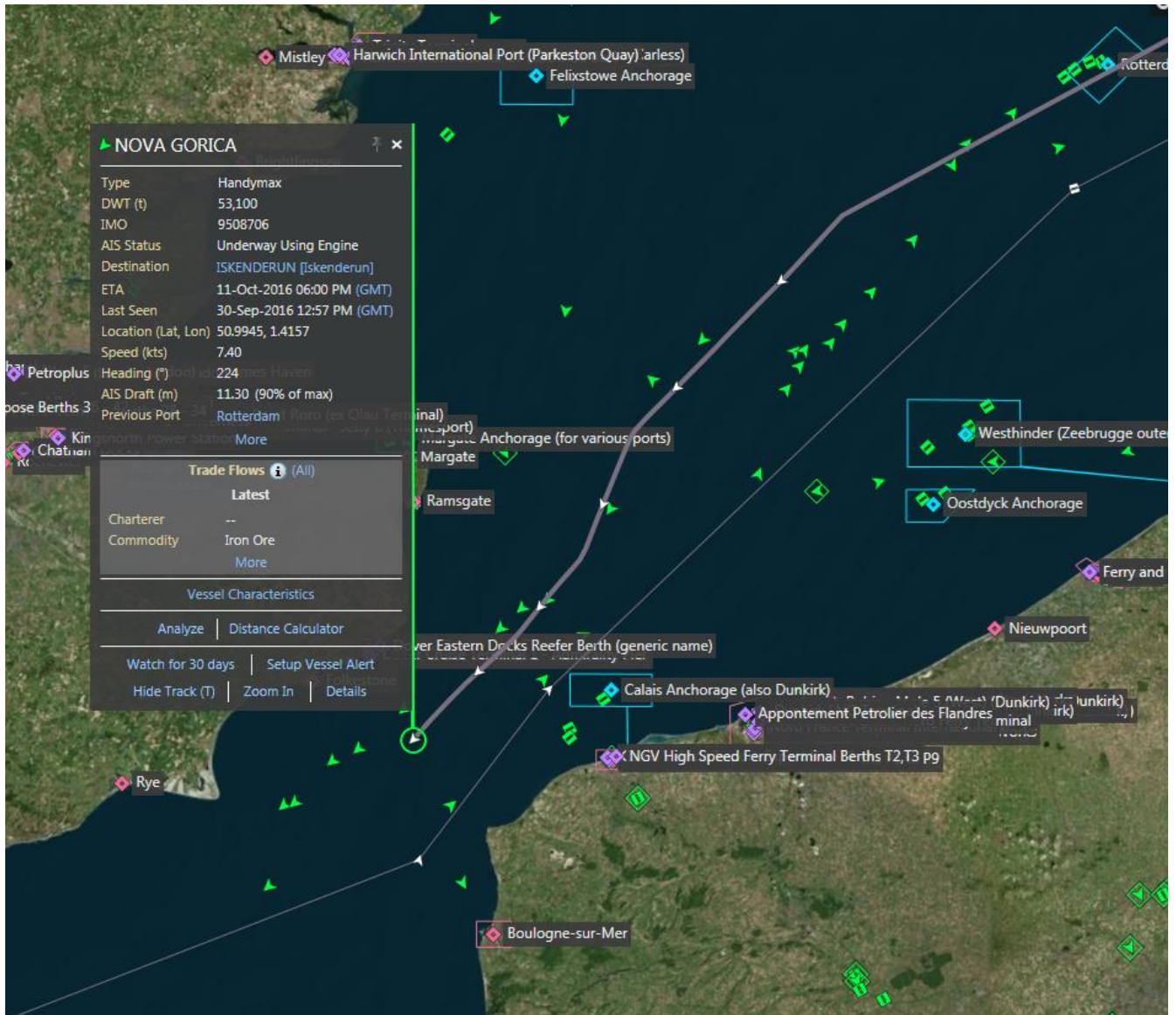
The majority of globally traded commodities (such as crude oil, coal and iron ore) are transported by sea from their country of production in large bulk carriers and tankers. For example, coal is mainly produced in Australia and South Africa and consumed in China. To avoid collisions, ships broadcast their locations and heading via a system called Automatic Identification System (AIS)[3]. While primarily intended for ship to ship communication, AIS data is also recorded and collected by land-based stations and satellite, providing a near-realtime view on the global transportation of many commodities and allowing changes in global flows to be observed over time. In addition to transportation data, storage data is also available for many markets allowing changes in flows to be tracked over time. When combined with existing supply/demand information, this data can provide a comprehensive overview of the underlying market dynamics.

TASK

The objective of the project is to construct a model for the transportation and storage of a set of global commodities based on information contained in AIS messages and to identify how changes in flow over time can impact commodity prices. This is likely to include both short-term effects (for example congestion at the Suez canal causing delays) and longer term effects (changes in rates of consumption across different countries). The work is expected to build on ongoing research at Cantab that uses recurrent neural networks to identify suitable relationships, although the student is encouraged to suggest and experiment with alternative approaches. The expected output of the project will be a library of code and modelling techniques with proven applicability to the applications of interest and a report summarising the project results.

SUGGESTED READING:

1. Ferrazzi et. al. Bayesian approaches to reverse engineer cellular systems: a simulation study on non-linear Gaussian networks. BMC Bioinformatics 2007.

2. Keras documentation: https://keras.io/

3. Automatic Identification System: http://www.allaboutais.com/

NOVA GORICA

Type	Handymax
DWT (t)	53,100
IMO	9508706
AIS Status	Underway Using Engine
Destination	ISKENDERUN [Iskenderun]
ETA	11-Oct-2016 06:00 PM (GMT)
Last Seen	30-Sep-2016 12:57 PM (GMT)
Location (Lat, Lon)	50.9945, 1.4157
Speed (kts)	7.40
Heading (°)	224
AIS Draft (m)	11.30 (90% of max)
Previous Port	Rotterdam
	More

Trade Flows ⓘ (All)

Latest

Charterer	--
Commodity	Iron Ore
	More

Vessel Characteristics

Analyze | Distance Calculator

Watch for 30 days | Setup Vessel Alert

Hide Track (T) | Zoom In | Details

Recommended engine implementation for the development of tailored and revenue-driven promotions
(Marco Manetti)

Thanks to the information stored in the company data warehouse, **Costa Crociere** collects all the information that characterize the ticket purchasing behavior, the demographic and the on-board experience of its customers. Such data, provided with the appropriate modeling algorithm, represent a key element for supporting the company decision making [4]. In more details, the analysis and the management of such big amount of information, allows not only the profiling of the company customers, but also the determination of their purchasing behavior [1,3]. Furthermore, the mining of historical booking data represents a powerful tool for the demand-forecasting of on-board services [2]. By matching all those different information, it is possible to develop tailored call-to-actions and promotion campaigns, ensuring the final revenue maximization by minimizing the dilution risk. While such a process is already well established when we treat loyal customers with a solid and rich profiling, the same analysis becomes more challenging when we treat with first times customers, from which only limited information are available. The proposed project is aimed at the development of a predictive engine for on board services trading, with the final output of the future customers purchasing behavior mapping, as a function of selected leverage parameters, which could provide a significant insight for the development of tailored promotions.

To sum up, the project will be divided in the following stages:
1. Analysis, cleaning and mining of the historical booking and on-board items selling data
2. Qualitative and quantitative detection of leverage parameters for different on board services purchase
3. Prediction of future passengers purchasing behavior, based on historical data modelling
4. Model testing and complete analysis of its potential and limits

References
1. Shaw, Michael J., et al. "Knowledge management and data mining for marketing." *Decision support systems* 31.1 (2001): 127-137.
2. Fildes, Robert, et al. "Forecasting and operational research: a review." *Journal of the Operational Research Society* 59.9 (2008): 1150-1172.
3. Ngai, Eric WT, Li Xiu, and Dorothy CK Chau. "Application of data mining techniques in customer relationship management: A literature review and classification." *Expert systems with applications* 36.2 (2009): 2592-2602.
4. Power, Daniel J., Ramesh Sharda, and Frada Burstein. "Decision support systems". John Wiley & Sons, Ltd, (2015).

Finding Structure, Symmetries and Dynamics in Graphs

Keith Hermiston, DSTL

A key component of future analysis capability will be the ability to handle vast datasets automatically, both at rest and streaming, using Big Data technology. A stressing mode of Big Data processing is that of graph analytics. Graph analytics is of most relevance where data sets comprise both entities and relationships. They typically occur in network datasets such as social-media, computer network traffic, spatial flow systems (e.g. transport networks), semantic models and natural language inference.

The project will focus a theme of future **urban analysis**, where large datasets capture the ebb and flow of normal population activity, often periodic. Within the datasets reside changing structures, symmetries and dynamics that provide a narrative of urban life at different levels of granularity. The main challenge of the project is to familiarize the researcher with graph analysis concepts, tools and methodology and to offer an introduction to graph analytics [1], with the opportunity to innovate.

The project will review the spectrum of graph research projects and datasets at the Stanford Network Analysis Project (SNAP) [2]. Selecting a topic of relevance to network analysis and graph mining of urban datasets, the project will apply a selection of SNAP code (perhaps using the Python libraries) to investigate (detect and characterise) anomalies, structures and periodic temporal patterns within graph datasets at different levels of granularity. Understanding of the analysis tools (strengths and weaknesses) will be acquired and innovations of further functionality building upon existing analysis tools are encouraged. In selecting datasets the project should consider those drawn from both physical space and cyber space, reasoning over the analysis results to draw conclusions of significance. A further open data set that could be considered under this work is available from Transport for London [3].

References

[1]   Algebraic Graph Theory, Springer, ISBN 0-387-95220-9, 2001.

[2]   Stanford Network Analysis Project, Internet URL https://snap.stanford.edu/ [accessed 9th August 2016]

[3]   Transport for London, Open Data Users, available from https://tfl.gov.uk/info-for/open-data-users/ [accessed 9th August 2016]

Attenuation correction in SPECT CT images

Marius DeGroot, GSK

Researchers at the University of Cambridge perform in-vivo imaging of neutrophils – particular white blood cells that form part of the immune system. For this procedure, neutrophils are isolated and labelled with technetium, after which they are (gently to avoid activation) reinjected in the bloodstream of the study participant.[1] Single photon emission computed tomography (SPECT) imaging can then be used to 'follow' the neutrophils, as they are recruited in a triggered immune response. The dynamics of these responses, mean that imaging needs to be acquired over multiple scan sessions during the course of half a day.

SPECT imaging, which is intrinsically limited in spatial resolution, is routinely collected in conjunction with a computed tomography (CT) image for enhanced anatomical understanding. Routinely, the CT is also used to correct the SPECT image intensities for location-specific differences in the attenuation that the photons experience depending on the volume and density of tissue that needs to be traversed before reaching the detectors.[2] However, CT scanning exposes the study participants to ionizing radiation. In order to reduce this exposure in the neutrophil imaging experiments, a CT image is only acquired with the first SPECT image in the experiment and not for the followup sessions recording the immune response. Because participants are repositioned in the scanner for each image, there is a correspondence problem between the baseline CT and later SPECT images, which can currently only be solved by taking the images outside of the clinical systems and aligning images using image registration. The aim of this project is therefore the formulation of an attenuation correction algorithm using the transformed CT images.

Photons emitted by the radioactive technetium undergo Compton scattering proportional to the radiodensity and length of tissue needed to be traversed before being able to exit the body. Energy loss and change or direction for scattered photons, means that the detection rate for more heavily affected photons is reduced. Reconstructing a hypothetical object with a uniform radioactivity, would result in an inhomogeneous appearance with a lower activity towards the centre of the object. Building a transmission based attenuation correction model for SPECT CT, needs to take into account the differential transmission physics of electrons at various energies (as the CT image uses photons with a different energy than the X-rays emitted by the technetium), and hinges on estimating (linear) correction factors for SPECT from the reconstructed CT image. Multiple approaches, of varying complexity, have been described to estimate these correction factors.[3]

Addition of attenuation correction to the neutrophil imaging workflow described above, is thought to enhance accuracy of quantification and will contribute to this active line of research.

References

1. Ruparelia, P. *et al.* Quantification of neutrophil migration into the lungs of patients with chronic obstructive pulmonary disease. *Eur. J. Nucl. Med. Mol. Imaging* **38,** 911–9 (2011).

2. Patton, J. A. & Turkington, T. G. SPECT/CT physical principles and attenuation correction. *J. Nucl. Med. Technol.* **36,** 1–10 (2008).

3. Zaidi, H. & Hasegawa, B. Determination of the attenuation map in emission tomography. *J. Nucl. Med.* **44,** 291–315 (2003).

Title: Model-based experimental design in imaging studies on animals: how much can we reduce animal numbers?

Background: In drug development we need robust statistics to build confidence around findings. Datasets are used to compute statistics based on the distribution of noise and signal which typically require of large number of animals in a factorial design [1]. However, pressures from regulators and ethical committees restrict the animal numbers to be researched upon [2].

Mechanistic models utilise knowledge of the biology to describe the data and simulate scenarios. Mechanistic knowledge includes information from nano to macro scales, such as drug-target interaction, drug chemistry and ADME[1] properties, intracellular pathways, tissue and cellular reorganisation, etc. These mechanisms can be mathematically described, as it is common practice in other industries such as engineering [3].

Imaging studies within drug development depend heavily on the parameterisation of the image analysis and its manual curation. This extra source of noise introduced in the data is difficult to characterise and often presents with skewed distributions [4]. The restricted animal numbers and the many sources of variability in imaging studies put at risk the validity of our experiments whereby statistical models often offer pitfalls, however there are mechanistic modelling alternatives that can improve our experimental designs.

Problem: Let Y be our set of measurable outputs and $\widetilde{Y}(\theta, x)$ be a general known function of the parameter set θ and state variables x. The likelihood ($L(\theta|x)$) of the optimisation problem can be approximated to a least-squares objective function $\min_{\theta \in \Omega} L(\theta|x) \approx \left(\widetilde{Y}(\theta, x) - Y\right)^2$, where the parameter sets (θ) can be chosen from the feasible parameter space Ω and subject to known system constraints. The sensitivity of the problem can be evaluated by calculating the Jacobian matrix (S) as the first derivative with respect of the parameters as $S(\theta, x) \approx \frac{\partial(\widetilde{Y}(\theta,x) - Y)}{\partial \theta}$. This definition allows us to approximate the Fisher Information Matrix (FIM, $\mathcal{J}(\theta, x)$) to $\mathcal{J}(\theta, x) \approx S^T \times \Sigma^{-1} \times S$ where Σ represents the variance-covariance of the data in matrix form. Fisher information is an important element in dynamic modelling to determine practical model identifiability and confidence in the parameter estimates. However, Σ represents a degree of expected noise in the data and its classical calculation relies on a significant amount of repeats or animal numbers [5].

Aim: This project will explore alternatives to factorial experimental design using imaging studies performed in small animals (mainly rats and mice) to determine key dose-response efficacy/toxicity endpoints within drug discovery. Concretely, we want to focus on the derivation of the previously-defined data noise matrix Σ to improve the amount of Fisher Information and therefore confidence in our model's parameter estimates with a special focus on reducing animal numbers. A theoretical framework will be initially developed, followed by validatory experiments based on the results of the theory.

Expectations: The student will mainly work with a non-linear multidimensional optimization problem as described above. She/he will be expected to explore different methods to estimate the noise matrix Σ and FIM, and design a large dose-response study for a concrete problem by the end of their

---

[1] ADME: Administration, Distribution, Metabolism and Elimination of a drug. These properties determine the life cycle of a drug in the body also known as pharmacokinetics.

stay. Additionally, she/he can (if required/desired) have an active input in designing routines of image analysis, registration, and reconstruction for a variety of imaging techniques. Further, she/he will have the opportunity to experience the practicalities of the different in-vivo imaging modalities, such as computed tomography (CT), magnetic resonance imaging (MRI) and optical (bioluminescence (BLI) and fluorescence).

Minimum requirements are a good understanding of basic statistics, Euclidean geometry, and ease with non-linear programming/optimisation methods [5]. This is an exciting opportunity for a motivated student to make a difference in the ways experiments are designed and conducted within the in-vivo bioimaging community at the world-leading pharmaceutical company GSK, impacting thereby the way we deliver safe and efficacious drugs to patients.

References:

1. Winer BJ, Brown DR, Michels KM: **Statistical principles in experimental design**, vol. 2: McGraw-Hill New York; 1971.
2. Webpage HO: **https://www.gov.uk/guidance/research-and-testing-using-animals**.
3. Franceschini G, Macchietto S: **Model-based design of experiments for parameter precision: State of the art**. *Chemical Engineering Science* 2008, **63**(19):4846-4872.
4. Duncan JS, Ayache N: **Medical image analysis: Progress over two decades and the challenges ahead**. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2000, **22**(1):85-106.
5. Bertsekas DP: **Nonlinear programming**: Athena scientific Belmont; 1999.

Multi-parametric Textural Analysis of Cartilage.

Josh Kaggie, GSK

Osteoarthritis (OA), a common joint disease, leads to reductions in the spatial organisation of cartilaginous structures. Textural analysis is a mathematical method that has been proposed to measure quantifiable image-related information associated with OA instead of subjective, reader-dependent information. Grey-level co-occurrence matrix (GLCM) mean, contrast, and variance measurements have shown significant increases in subjects at risk for OA [1,2]. Textural analysis can lead to improved visualization and characterization of disease, improving the ability to accurately and repeatedly measure cartilage thickness.

Medical imaging techniques, such as Magnetic Resonance Imaging (MRI), give multi-parametric information, such as three spatial dimensions and one time dimension. While there are commercial solutions for textural analysis in two and three dimensions, the implementations above two dimensions in medical imaging are rare. This 3-month CDT call is for a student to implement textural analysis on multi-parametric MRI knee data.

[1] Joseph, Gabby B., et al. "Texture analysis of cartilage T 2 maps: individuals with risk factors for OA have higher and more heterogeneous knee cartilage MR T 2 compared to normal controls-data from the osteoarthritis initiative." Arthritis research & therapy 13.5 (2011): 1.

[2] Boniatis, I. S., et al. "Quantitative assessment of hip osteoarthritis based on image texture analysis." The British journal of radiology (2014).

A source estimation inverse problem in oilfield drilling: Correcting surface arrival times of drilled cuttings for the effects of dispersion during transport up the well

Background

When a hydrocarbon well is being drilled the rock cuttings created at the drill bit are hydraulically conveyed to surface for disposal by flow of drilling fluid ("mud") upwards in the annulus between the rotating drill pipe and the rock. In many cases, measurements made on cuttings retrieved at surface ("mud logging") are the most economical way to obtain information about the formations being drilled, this information then being used to guide decisions on the locations of productive strata and on when to stop drilling. The position of the drill bit at any instant is known, as is the flow rate of the mud, and so it might seem a trivial matter to relate the time of arrival of a particular cutting at surface with the depth at which it was created, however this ignores the fact that transport of cuttings up the annulus is not a simple conveyor belt process but is subject to significant axial dispersion arising from a variety of physical processes.

The central aim of this project is to devise and assess mathematical algorithms for correcting time histories of cuttings properties measured at surface for the effects of dispersion during transport so as to create a map of the varying formation properties versus depth. More formally, we seek to solve a source estimation inverse problem for the one-dimensional advection-diffusion equation given boundary data and a known source location, so as to determine the properties of the downhole rock formations which give rise to the observed surface measurements as a consequence of transport up the well by an only partly-characterized convection-diffusion process.

Problem

The following two figures illustrate the effects of dispersion on the transport of small cuttings, with Figure 1 showing what we would like, and Figure 2 showing what we are likely to get in reality. The first of these takes a small value of axial dispersion coefficient (assumed constant), and the second a large value. The concentration (mass per unit volume) of formation material of species i at the exit of the well, $W_i(0,t)$, was computed from an analytical solution of the advection diffusion equation

$$\frac{\partial W_i}{\partial t} \quad V\frac{\partial W_i}{\partial x} \quad D\frac{\partial^2 W_i}{\partial x^2} = UW_i^{rock}\big(L(t)\big)\delta\big(x \quad L(t)\big), \tag{1}$$

obtained through linear superposition in the form

$$W_i(0,t) = \int_0^t UW_i^{rock}\big(L(t')\big)\frac{\exp\Big(\quad \big(V(t \quad t') \quad L(t')\big)^2/\big(4D(t \quad t')\big)\Big)}{\sqrt{4\pi D(t \quad t')}}\,dt', \tag{2}$$

where U is the (dimensionless) rate of penetration (assumed constant in time), $L(t) = L_0 + Ut$ is the (dimensionless) depth of the drill bit and location of the source of the cuttings, V is the (dimensionless) mud circulation velocity (assumed constant in time), W_i^{rock} is the composition of the formation, and D is the (dimensionless) coefficient of axial dispersion/diffusion (assumed

constant). It is straightforward in principle to replace the analytical solution (2) with a numerical solution taking account of non-constant annulus cross sectional area, of time-varying values of U and V, and using more realistic values for D e.g. making it dependent on V so as to properly represent Taylor dispersion. In the figures compositions measured at surface are lagged to downhole locations according to

$$\text{depth of provenance of cuttings emerging at } t, L(t) = \frac{V}{U+V}(Ut + L_0)$$

(3)

and compositions are corrected for dilution effects using

$$W_i^{rock}(L(t)) = \frac{U+V}{U}W_i(0,t) \ .$$

(4)

These expressions assume that the mud flow rates, and rate of penetration, are constant in time, but could be generalized. We note in passing that if the $L(t')$ term in the exponential term in equation (2) were a constant then that expression would be a standard convolution operation with a translationally invariant kernel, however in actuality the kernel varies from point to point, and the time dependence of $L(t')$ is a contributor to the precise arrival times of material at surface, so (2) must be thought of as $W_i(0,t) = \int_0^t W_i^{rock}(t')f(t \quad t',t')dt'$ and not $\int_0^t W_i^{rock}(t')f(t \quad t')dt'$.
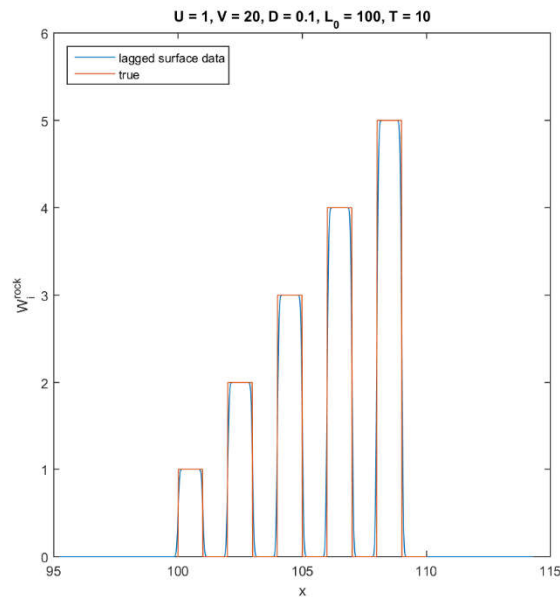

Figure 1: Cuttings transport is simulated with and advection-diffusion equation. The simulated concentrations are then lagged, and scaled to correct for dilution, to as to estimate the incoming concentration from the formation. For this example the (dimensionless) diffusion coefficient takes the value D=0.1, and the effects of hydrodynamic dispersion are small. The (dimensionless) rate of penetration U=1 between (dimensionless) times 0 and T=10 and zero thereafter, the mud circulation velocity is V=10, the initial depth of the well is L₀=100. The data is simulated out to $t = 20$. The estimated concentrations are close to the true values, which correctly located in depth by simple lagging.
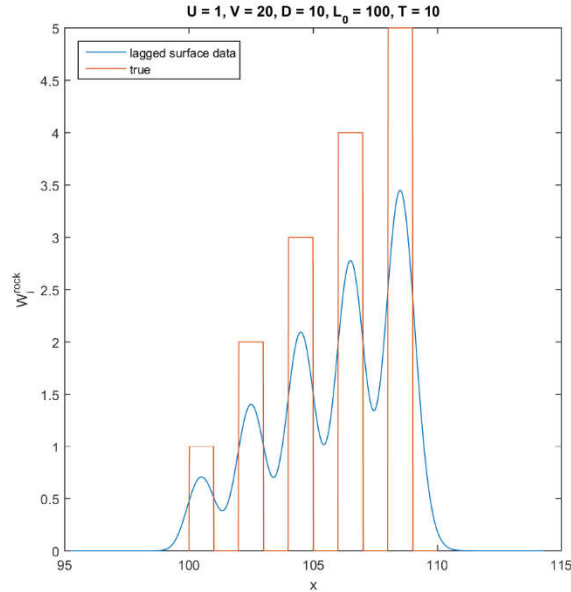
Figure 2: As previous figure, but now the (dimensionless) diffusion coefficient D=10. Note the smearing of spatial structure in the estimated formation concentration, and the numerical discrepancies between the true and estimated formation concentrations. Measurements errors of various kinds will also occur in reality, but are not simulated here.

In this project we seek a practical mathematical process for converting the observed cuttings composition data, $W_i(0, t)$, to a downhole log of composition, $W_i^{rock}(x)$ (the suffix i indicates that we might have several independent channels of data, e.g. concentrations of different chemical species, and this might perhaps be exploited to improve the quality of the estimation process). Solving such an inverse problem is known to be an ill-posed mathematical operation in general.

One approach is to minimize a suitable norm of the desired output, say $\left\| W_i^{rock}(x) \right\|_1$ subject to non-negativity, $W_i^{rock}(x) \geq 0$, and $\left\| W_i(0, t_j) \quad \int_0^{t_j} W_i^{rock}(L(t')) f(t_j \quad t', t') dt' \right\|_1 \leq \sigma$, which expresses consistency with the M observations, σ being an estimate of the error levels in the data, and f being the kernel function appearing in equation (2). (Note use of the 1-norm, which in crude numerical experiments seems to give better results than the 2-norm.) The parameters $\{U, V, D\}$ of G may initially be treated as known, but eventually D at least will have to be estimated as part of the process. Results applying this approach to the data of Figure 2, assuming the flow parameters are known, are shown in Figure 3. The algorithm successfully sharpens the shape of the formation composition, allowing a better depiction of the layering and giving reasonable values for the layer properties, yet is surely far from optimal and it is not informed by judicious use of the literature (some possible starting points are listed in the Useful Reading section below).

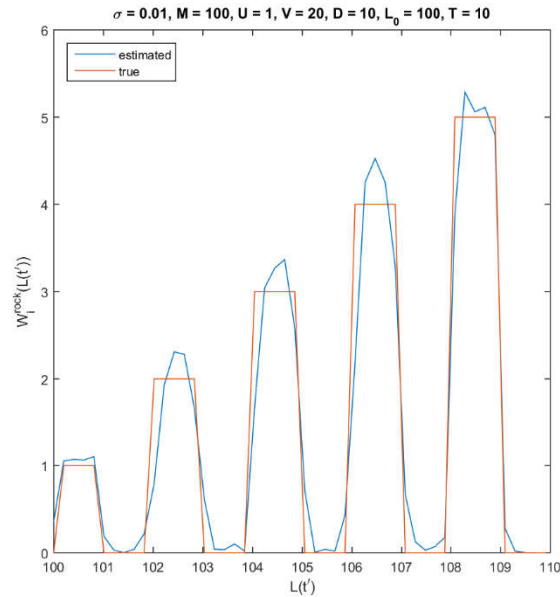$\sigma = 0.01$, M = 100, U = 1, V = 20, D = 10, L_0 = 100, T = 10

Figure 3: Result of estimating the rock properties by correcting the data of Figure 2 for the effects of diffusion using the minimization formulation. The formation composition is represented by values at 50 equispaced points, and the minimization is performed using an active set algorithm which tests show is more effective than the others available in the Matlab function fmincon.

Scope

Formulate the formation properties estimation problem appropriately (e.g. in Bayesian terms, as a source estimation problem, …). Devise one or more algorithms for the estimation of $W_i^{rock}(x)$, and explore their performance using synthetic data treating $\{U, V, D\}$ as known and including measurement errors. How little measured data can we get away with? Extend to case where there is uncertainty on the values of $\{U, V, D\}$, D being the parameter for which uncertainty is greatest and the poorest prior information will be available. Write up, and present conclusions to sponsor.

Should time and inclinations permit, there is certainly scope to look carefully at the formulation of the underlying transport model, to incorporate better representations of the physical processes involved in transport of dense particles in the annulus while drilling, and thence to improve our priors for D. A further interesting but separate question is the classification of measured cuttings compositions, or estimated formation compositions, into groups having similar properties.

Project aims & expectations

The main desired outcomes are: a mathematically respectable formulation of the formation composition estimation problem; a listing of potentially useable approaches and algorithms, with bibliography; results from applying selected algorithms on synthetic data; a written report detailing all the above. By the end of the project we would like to have received sufficient information, through the report and through conversations over the course of the activity, that we could implement some of the algorithms ourselves.

We are happy to host the candidate for a period in our research centre in Cambridge during the course of the project, duration and details to be arranged.

Useful reading

Hydrocarbon well drilling:

1. http://www.slb.com/resources/oilfield_review/~/media/Files/resources/oilfield_review/defining_series/Defining-Drilling.ashx
2. http://www.slb.com/resources/oilfield_review/~/media/Files/resources/oilfield_review/defining_series/Defining-Directional-Drilling.ashx
3. http://www.slb.com/resources/oilfield_review/~/media/Files/resources/oilfield_review/defining_series/Defining-Drilling-Fluids.ashx

Mud logging:

1. http://www.slb.com/resources/oilfield_review/~/media/Files/resources/oilfield_review/defining_series/Defining-Mud-Logging.ashx
2. http://petrowiki.org/Cuttings_analysis_during_mud_logging
3. http://petrowiki.org/Formation_evaluation_during_mud_logging

Hydrodynamic dispersion:

1. Taylor, G. I. (1953). "Dispersion of soluble matter in fluid flowing slowly through a tube." Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **219**: 186-203.
2. Taylor, G. (1954). "The Dispersion of Matter in Turbulent Flow through a Pipe." Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **223**(1155): 446-468.

Source estimation problem:

1. https://en.wikipedia.org/wiki/Inverse_problem
2. Badia, A. E., et al. (2005). "Identification of a point source in a linear advection–dispersion–reaction equation: application to a pollution source problem." Inverse Problems **21**: 1-17.
3. Mazaheri, M., et al. (2015). "Mathematical Model for Pollution Source Identification in Rivers." Environmental Forensics **16**(4): 310-321.   {contains useful overview & reference list}
4. Boano, F., et al. (2005). "Source identification in river pollution problems: A geostatistical approach." Water Resources Research **41**(7): W07023.   {a different approach}

Contact

Dr Paul S. Hammond, Schlumberger Gould Research, High Cross, Madingley Road, Cambridge CB3 0EL. pshammond@slb.com. +44 1223 325244.

02 September 2016

Project: Discrete Fourier transform over regular triangle or tetrahedron
Supervisor: Evren Yarman (Schlumberger)
Internal Contact: Carola-Bibiane Schönlieb (DAMTP)

In practice, measurements often have compactly supported Fourier transform, either by nature or due to the acquisition systems. For the interval 1D and tensor products of it, discrete Fourier transform is an indispensable analysis tool thanks to the development of fast Fourier transform (FFT), one of the most influaential algorithms of 20th century [1]. Considering a tesselation approximation of the support of the Fourier transform using simplexes, it is natural to explore extensions of discrete Fourier transform over 1D interval to simplexes, which in 2D and 3D are triangle and tetrahedron, respectively.

This project aims to extend the ideas of FFT to regular triangular and tetrahedral domains. In this regard, we would like to utilize numerical quadratures (Figure 1a) tailored to approximate the inverse Fourier transform of the characteristic function (Figures 1b and 2) over the given simplex and explore iterative recursive refinement combined with fast multipole method [4]. For 2D, one potential recursive refinement is a structure similar to the fractal Sierpinski's gasket [4] (Figure 1a). We will explore other refinements and their higher dimensional analogues. Applications areas include numerical analysis of images, videos and wavefields.

References

[1] Jack Dongarra, Francis Sullivan, "Guest Editors' Introduction: The Top 10 Algorithms", Computing in Science & Engineering, vol.2, no. 1, pp. 22-23, January/February 2000, doi:10.1109/MCISE.2000.814652

[2] Alan Edelman, Peter McCorquodale, and Sivan Toledo. The future fast Fourier transform? SIAM Journal on Scientific Computing, 20(3):1094–1114, 1998.

[3] Leslie Greengard and June-Yub Lee. Accelerating the nonuniform fast Fourier transform. SIAM review, 46(3):443–454, 2004.

[4] A. Lakhtakia et al. On the spatial Fourier transforms of the Pascal-Sierpinski gaskets. Journal of Physics A: Mathematical and General , 19: 3147-3152, 1986.
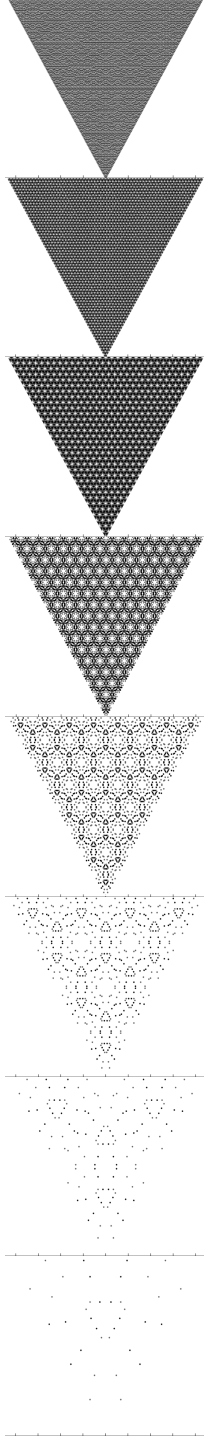
(a) Different quadratures used to approximate K in Figure 2.
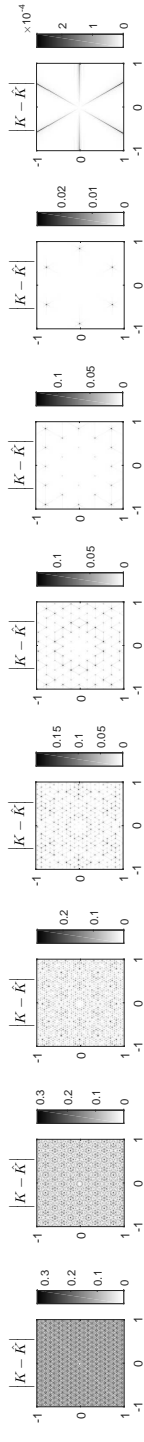
(b) Absolute error between the discretized Fourier approximation $\hat{K}$ of K and K, for quadratures above, respectively.

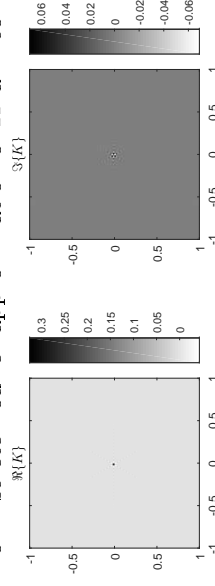Figure 1: Quadratures used for discrete Fourier approximation of K and corresponding absolute errors.

Figure 2: Real and imaginary parts of K, Fourier transform of the characteristic function over the triangle.

Fluid-Fluid Displacement in Boreholes

Simon Bittleston

September 27, 2016

1   Background

When drilling oil & gas wells it is necessary to make the borehole secure by inserting a steel tube (casing) into the well and filling the annular gap, between the borehole wall and the casing, with cement. Prior to a cementing job, the well is full of a non-Newtonian fluid know as drilling mud, which is a water or oil based fluid containing additives to achieve fluid properties which are required to satisfy operational criteria. After running the casing into the well, the placement of cement in the annulus is achieved by pumping it from the surface to the bottom of the well inside the casing and then back up the annulus. During the downward displacement, inside the casing, the flowing cement is segregated from the drilling mud by a bung (like a pipeline pig) which fits snugly within the casing. When the bung reaches the bottom the pressure at the surface is increased to break it, allowing cement to pass around the bottom of the casing (casing shoe) and up the annulus. As the cement is pumped up the annular space, it displaces the mud. This fluid-fluid interface is not completely flat, and there are a number of interesting instabilities which can take place. Only sufficient cement is pumped into the top of the casing to fill the annular space. It is followed by a second bung and more drilling mud so that by the end of the operation the inside of the casing is again filled with drilling mud. The diagram at the end of this document shows the sequence of events. (Although the diagram shows a straight vertical well, in reality most wells today have curved trajectories with horizontal sections within the reservoir.)

The reason for segregating the fluids on their downward journey inside the

casing is that the cement is generally heavier than the mud, which would lead to an unstable interface, and an extended mixing zone.

To make matters more complicated, an extra fluid is generally pumped between the mud and cement. It is called a spacer. The spacer fluid performs other functions one of which is to avoid direct contact between mud and cement. Often, mixing these can create a fluid with undesirable rheological properties.

Achieving a good cement job (i.e. a uniform sheath of cement between the casing and rock) is vital for ensuring zonal isolation and avoiding migration pathways in the annulus.

2   Problem

The annular displacement problem is now described more precisely. The key drilling mud properties are its density and rheology. A Herschel-Bulkley representation is generally sufficient for the mud, spacer and cement - although the specific characteristics of each can be quite different. All three fluids exhibit a yield stress, and variations of the rheologies between the three can strongly affect the motion of the displacement interface. The constitutive law for a Herchel-Bulkley fluid is:

$$D = 0 \qquad \text{for } |\tau| \leq \tau_y$$
$$\tau = (k\dot{\gamma}^{n-1} + \frac{\tau_y}{\dot{\gamma}})D \qquad \text{for } |\tau| > \tau_y$$

where D is the rate of strain tensor and $\dot{\gamma}$ its second invarient with τ is the stress tensor. The fluid properties are: k the consistency index, n the power index, and τ_y the yeild stress.

The casing is rarely precisely concentric in the borehole, although metal springs called centralisers are placed at intervals along the outside of the casing to push it towards the centre. Eccentricity of the casing is one of the key factors affecting the displacement process. The well is generally not vertical and may have sections which are horizontal. Since the mud, spacer and cement all have different densities, this again leads to variations in the interface between the fluids through gravity driven processes (e.g gravity

currents). The cross-sectional shape of the borehole is not uniform, leading to axial variations in its diameter. In some operations the casing may be rotated or oscillated up and down (reciprocated).

The displacement process may be performed at sufficiently low flow rates that all fluids are everywhere in laminar flow. In other instances, the rates may be high enough for all three fluids to be turbulent. Unfortunately, rates between these two can be required, leading to problems where individual fluids may be laminar, transitional or turbulent in different places; the fluids may have very different Reynolds numbers. A curious fact is that even for a single phase flow of a Newtonian fluid in an eccentric annulus (e.g. water) the fluid may be laminar on the narrow side of the annulus, turbulent on the wide side, and in transition between the two.

The problem is: given a well geometry, sequence of fluids, prescribed inlet flow rate (which may vary in time), calculate the displacement of the mud by the spacer and the spacer by the cement, in this three dimensional inclined eccentric annulus - for laminar, transitional and turbulent flow.

This problem has been studied before. There are useful small parameters which can be used to reduce the complexity, with some rough assumptions (not always satisfied), leading to a model system of equations. One such 'relatively' small parameter is the ratio of the annular gap to the circumference of the borehole. Expanding with a regular perturbation in this parameter, and taking the leading order terms, provides a tractable 2-D time dependent system of equations (SB(1)). One rough assumption is required to close the system of equations, which is that interface details across the annular gap can be replaced by a gap average.

The 2D model for the displacement process just mentioned is of the form:

$$\frac{DC}{Dt} = 0$$
$$\nabla.(f(C, |\nabla\psi|)\nabla\psi) = F(C, \theta, z, t, g)$$
$$(u, v)^T = (\frac{\partial\psi}{\partial z}, -\frac{\partial\psi}{\partial\theta})^T$$

where $u = u(\theta, z, t), v = v(\theta, z, t)$ are the average (across the annular gap) velocities in the azimuthal and axial directions, F is a forcing function which incorporates the effects of gravity, $C = C(\theta, z, t)$ is a tracer (level set) type function which tracks which fluid is present at each point in the domain at each time and f is a non-linear function which needs to be evaluated either

by an iterative scheme (at each point in the domain, at each time step and at each iteration of the solver for the elliptic equation!) or which, with certain assumptions, can be reduced to a look-up table in 5+ dimensional space (precomputed). f also depends on the rheology of the fluids.

In laminar flow function f is basically a ratio of pressure-drop to flow rate locally - this is a lubrication approximation. Similarly, assuming such a local view can be assumed, a turbulent flow relationship can be inserted when the local (gap) Reynolds number is sufficiently high. In this way the model equations can be adapted to predict displacements where parts of the flow are laminar and parts are turblent. Clearly this is a 'rough' approximation.

A second useful small parameter is the circumference of the well-bore divided by its length. For certain relatively simple conditions, one can approximate the whole displacement by a Taylor limit (RS(2)). In general such an approximation is not justified, e.g. in approximating unstable interfaces.

3 Scope

The flow of single phase Bingham plastics in complex geometries has received considerable attention in recent years. Early numerical solutions to simple problems regularised the equations by smoothing the stress-strain relationship for low shear rates. Various constructions were proposed for the non-linear stress strain relationship, but all exhibit difficulties in accurately determining the yield surface between shearing regions and plug flow (or no flow) regions. At the boundary the curvature of the velocity field is discontinuous. This creates low order convergence for numerical approximations to the elliptic operator.

Recent advances in convex analysis have provided new numerical methods for flows with yield stresses. Perhaps the best paper is JR(3). The awkward rheological problem of having a yield stress is included through a dual constraint - duality methods are combined with semi-smooth Newton methods to provide a fast and robust solver.

The 2-D time dependent approximate model equations for the displacement problem exhibit a similar difficulty to the single phase yield stress fluid as a sufficiently high local pressure gradient is required to mobilize the fluid

and below that there is no flow – this strong non-linearity is the root of the difficulties with the problem. The first part of the work will be to to apply these recent convex methods to the model displacement equations. This ought to lead to fast solvers for that system of equations.

It may also be possible to apply matched asymptotic expansions to solve the non-linear elliptic equation to provide sufficiently accurate approximations, or at least approximations which can be used as a worthy starting point for a fixed point iteration.

Further a detailed reanalysis of the full 3D-time dependent, multi-fluid problem may produce a more representative system of model equations, or perhaps, a coupled system which represents the details of the flow across the annular gap more accurately.

4   Project aims & expectations

There are several possible directions this work could take. Amongst these are:

1. Apply convex analysis and duality methods to create a fast solver for the non-linear elliptic equation for the model equations for fluid-fluid displacement

2. Show, theoretically, that the resulting method converges, derive the algorithm convergence rate, and determine robustness

3. Investigate higher order approximations for problems which have discontinuous curvature, particular those where boundaries, like yield surfaces, must be determined as part of the solution procedure

As a follow on Ph.D. project this can be extended to include:

1. Attempt to produce a matched asymptotic expansion for the fully non-linear elliptic problem, or simpler versions of this equation which arise when assumptions are relaxed

2. Revisit the full 3D system to seek more complete approximate models which better represent across-annular-gap instabilities

3. Compare model results with available experimental measurements

Overall, it should be emphasised that a fast, robust solver which captures the important aspects of this displacement problem is highly desirable, since cementing operations are performed on every oil well drilled across the world.
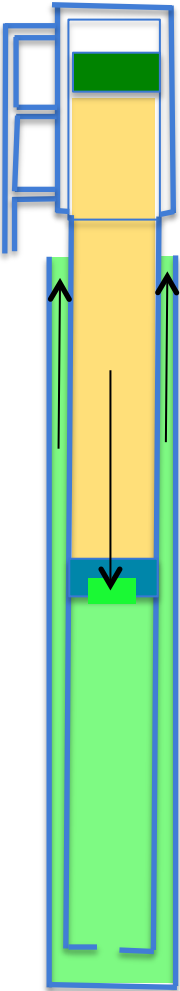
5   References

SB(1) Mud removal and cement placement during primary cementing of an oil well Laminar non-Newtonian displacements in an eccentric annular Hele-Shaw cell. Journal of Engineering Mathematics (2002). S.H. Bittleston, J. Ferguson, I.A. Frigaard

RS(2) Two-dimensional shear dispersion for skewed flows in narrow gaps between moving surfaces. Journal of Fluid Mechanics 214(-1):211 - 228 May 1990, Ronald Smith

JR(3) A duality based semismooth Newton framework for solving variational inequalities of the second kind. Interfaces and Free Boundaries 13 (2011). Jab Carlos de Los Reyes, Michael Hintermuller.
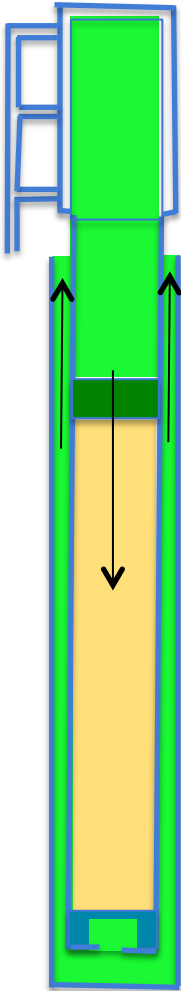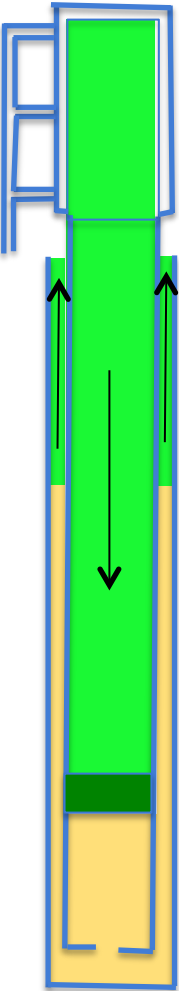
Cementing sequence
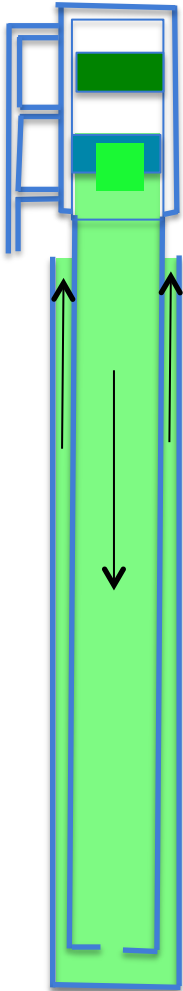
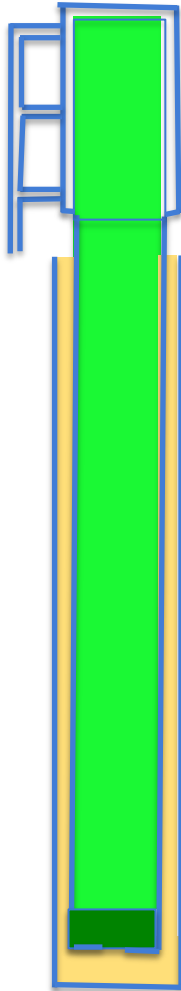Mud Circulation Cement inside Casing Bump Plug Displacement Finished

Fully Bayesian analysis of multi-exponential decays in low resolution NMR

Background

Nuclear Magnetic Resonance (NMR) is a premium measurement physics in petroleum well-logging [1]. Petrophysical NMR is typically restricted to time-domain data, analysed for distributions of the NMR relaxation times T_1 and T_2, and diffusivity. In the laboratory, analysis for distributions of pore sizes, induced internal fields and molecular motion correlation times may be added. Interpretations are usually based on integrals over finite ranges, other moments, or marginals. The same measurements have application to other porous media e.g. catalysts, cement-based materials, soil science etc [2]. Analysis of discretely-sampled time-domain signals in terms of multi-exponential decays with a broad distribution of relaxation times is ubiquitous. Experiments may be extended to multiple time or diffusion-encoding dimensions, analysed for joint distributions of two or more NMR parameters. Exponential decays are far from unique to NMR, but multiple dimensions may be unusual.

These NMR applications have enjoyed great success in multiple contexts. However, a major shortcoming of current routine practices, unresolved in 25 years, is that no error bars or other uncertainty metrics are available; nor are there easy ways to impose known physical constraints e.g. upper limits on relaxation times, or $T_2 \leq T_1$. Computational demands make some approaches impractical in dimensions higher than 2, and in 2D experiments negative amplitudes are known to be possible, though unconditionally disallowed by current analysis codes.

The archetype problem is the solution, for an unknown distribution $P(T_2, T_1) \geq 0$, of a First Kind Fredholm integral equation with exponential kernel

$$M\left(\tau_1^{(i)}, \tau_2^{(j)}\right) = \int_0^\infty P(T_2, T_1) \exp\left(-\tau_2^{(j)}/T_2\right) \left[1 - \exp\left(-\tau_1^{(i)}/T_1\right)\right] dT_2 dT_1 + e\left(\tau_1^{(i)}, \tau_2^{(j)}\right)$$

where $M\left(\tau_1^{(i)}, \tau_2^{(j)}\right)$ represents experimental magnetization values at discrete sampling times $\tau_1^{(i)}, \tau_2^{(j)}$ for $i = 1, \ldots N; j = 1, \ldots M$; with $e\left(\tau_1^{(i)}, \tau_2^{(j)}\right)$ being experimental noise (usually Gaussian, and white). Signal to Noise Ratio may be poor. Different measurements (diffusion, molecular correlation times etc) may involve kernels of more complicated form. The kernel may or may not be separable $K(\tau_1, \tau_2; T_1, T_2) = k_1(\tau_1; T_1) k_2(\tau_2; T_2)$, and the condition $P(T_2, T_1) \geq 0$ may need to be relaxed.

Current routine practice employs one of many *ad hoc* variants of Tikhonov regularisation [3,4,5,6], with new flavours still being regularly produced e.g.[7]. Attempts to assign uncertainty estimates include: (i) confidence intervals on the overall data misfit (with explicit rejection of Bayesian priors) [8], (ii) direct estimation of petrophysically important moments, via the use of Mellin transforms [9, 10], and (iii) Markov Chain Monte-Carlo (MC²) exploration of posterior probability spaces, but without close attention to the chosen priors [11].

Systematic treatment as a problem in Bayesian inference does not appear to have been attempted (in this application context). Nevertheless the apparatus for doing so may have been long available. Independence of discretization imposes constraints on the functional form of a consistent Bayesian prior [12]. Random samples from a wide class of uncorrelated measures, appropriate as priors, are known to be "atomic" (i.e. encodable as a finite set of discrete masses) [12,13,14]. The "atomic" property is arguably reasonable in the context (diffusion of liquid molecules to relaxing surfaces in porous media); and has also been made the basis of practical implementations [15] with various "engines" for MC² exploration of posterior probability spaces. There is demonstrated potential for resolving particularly challenging test cases and concomitant estimates of the Bayesian "evidence" ("prior predictive") are also available.

The Project

This project proposes two lines of enquiry:

1. Use of the public domain (LGPL) programs BayeSys and MassInf [15] , or other suitable tools, to explore the estimation of unknown distributions $P(T_2)$ (and higher dimensional generalizations) given experimental data records, with benchmarking of performance against current methods. Establishing practical computational limits (processing and memory) of the proposed approach will be an important overall objective, for evaluation against current practice.

2. Comparison of the theoretical basis of existing methods with the systematic Bayesian approach proposed. This should inform approaches to the visual display of results, to make these assimilable by users such as the practising petrophysicist. Scope for merit-ranking of underlying physical models can also be explored.

Experimental data records are available in profusion and test cases under controls can readily be acquired in the low-field NMR laboratory at SGR. Field data records (well logs i.e. multiple depth cases of the basic problem) could also be provided. Numerical test models used in the development of other approaches can be provided for the generation of synthetic test data. Analysis codes used in SGR and elsewhere in current approaches are available for benchmarking.

The overall scope of the project will ultimately be determined by the time commitments of the student. However the central objective of using publicly-available software to demonstrate the viability (or not ?) of a fully Bayesian approach on the archetype one-dimensional case, should be achievable within the stated timeline.

References

[1] W E Kenyon (1997). Petrophysical principles of applications of NMR Logging. *The Log Analyst,* March-April 1997, 21-43

[2] J Mitchell, L F Gladden, T C Chandrasekera & E J Fordham (2014). Low-field permanent magnets for industrial process and quality control. *Prog Nucl Mag Res Spectros*, **76**, 1-60

[3] A N Tikhonov & V Y Arsenin (1977). *Solutions of ill-posed problems.* London, Wiley.

[4] E J Fordham, A Sezginer & L D Hall (1995). Imaging multiexponential relaxation in the $(y, \log T_1)$ plane, with application to clay filtration in rock cores. *J Mag Res Ser A* **113** 139-150

[5] L Venkataramanan, Y-Q Song & M D Hürlimann (2002). Solving Fredholm integrals of the first kind with tensor product structure in 2 and 2.5 dimensions. *IEEE Trans Signal Process* **50** 1017–1026

[6] J Mitchell, T Chandrasekera & L F Gladden (2012). Numerical estimation of relaxation and diffusion distributions in two dimensions. *Prog Nucl Mag Res Spectros* **62** 34–50

[7] V Bortolotti, R J S Brown, P Fantazzini, G Landi & F. Zama (2016). Improved inversion of two-dimensional NMR relaxation data with the UPEN principle. *Proc. 13th Int. Conf. on Magnetic Resonance in Porous Media, Bologna,* Sept 2016.

[8] R L Parker & Y-Q Song (2005). Assigning uncertainties in the inversion of NMR relaxation data. *J Mag Res* **174** 314-324

[9] L Venkataramanan T M Habashy D E Freed & F K Gruber (2012). Continuous moment estimation of CPMG data using Mellin transform. *J Mag Res* **216** 43-52

[10] F K Gruber, L Venkataramanan, T M Habashy & D E Freed (2013). Estimation of petrophysical and fluid properties using integral transforms in NMR. *J Mag Res* **228** 104–115

[11] M Prange & Y-Q Song (2010). Understanding NMR T_2 spectral uncertainty. *J Mag Res* **204** 118–123

[12] S Sibisi & J Skilling (1997). Prior distributions on measure space. *J Roy Stat Soc B* **59**(1) 217-235

[13] T S Ferguson (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statistics* **1**(2) 209-230

[14] J F C Kingman (1993). *Poisson Processes,* OUP.

[15] Skilling, J. (2004) *BayeSys and MassInf* (2004). Manual for the eponymous softwares. Trademarked and Copyright by Maximum Entropy Data Consultants Ltd. Distributed in the public domain under the GNU Lesser General Public License.