2018 INDUSTRIAL PROBLEM SOLVING WORKSHOP (IPSW)

UPDATED: New Room Numbers!

August 19-24, 2018
The University of Calgary

THE INSTITUTES' INDUSTRIAL PROBLEM SOLVING WORKSHOPS aim to create mutually beneficial links between industrial researchers and their counterparts in academia. The goal of the IPSW is to connect industries with faculty, postdoctoral fellows and graduate students with expertise in industrial case-studies. This interaction is fostered in the specific context of a problem-solving session over 5 days. The case-studies in question have a significant mathematical or statistical content and in many instances lead to further collaboration and industrial development.

ORGANIZERS:

Michael Lamoureux (UofC); Odile Marcotte (UQÀM); Tom Salisbury (York); Cristian Rios (UofC)

SCHEDULE:

SUNDAY, AUGUST 19
Morning: High-performance computing bootcamp
Evening: Ice-breaker reception

MONDAY, AUGUST 20:
Morning: Presentation of Problems
Afternoon: Team break-out, problem study

TUESDAY TO THURSDAY, AUGUST 21-23:
All Day: Examine problems and report solutions (Possible excursion one afternoon)

FRIDAY, AUGUST 24:
Morning: Report Summary Presentations from teams

For more information, please visit:

http://www.pims.math.ca/industrial-event/180819-ipsw

Pacific Institute for the Mathematical Sciences

FIELDS

NSERC CRSNG

www.pims.math.ca @pimsmath facebook.com/pimsmath www.mathtube.org

Getting Started

Get connected: Select the "airuc-guest" wireless network on your wireless device. Open up a web browser, and you will be directed to the login page.

You will receive password via email. Info: https://conted.ucalgary.ca/info/wifi.jsp

Or use the Eduroam network, if you have an active Eduroam account.

Frequently Asked Questions

Q: Where is my hotel?

Most out-of-town participants have booked rooms at Hotel Alma, 169 University Gate NW.

Q: What's happening Sunday?

There is an optional "ice breaker" at Jameson's Pub, north of campus. Drinks and appetizers, and a chance to meet with fellow participants.

Q: Where do I register for the conference on the first day?

Check-in and package pick up can be done Monday morning in Science Theatre 141 (the Foyer)

Q: Where are the sessions?

All sessions are on main campus, University of Calgary. Joint sessions are in the Science Theatres ST 141, team meetings in Math Science building, MS 325, MS 365, MS 427, MS 452, MS 569. Computer Lab in MS 317. Coffee in ST 141 Foyer, or MS 371.

Q: Where can I access computers with mathematical software?

Via a web browser, access https://pims.syzygy.ca/ for Python, R tools. MS 317 has computers.

Q: Where can I go for help on site?

If you need assistance or have a question during the conference, please feel free to talk to one of the organizers. The Math & Stat Department can also help, at 403-220-3951.

Q: Where can I get refreshments and meals?

The MacEwan Student Centre has food and amenities:

https://www.su.ucalgary.ca/about/who-we-are/macewan-student-centre/

The workshop will provide morning coffee and pastries and afternoon coffee breaks

Q: Where can I get directions for campus and the building?

You will find a campus map on page 3, and a building map of the conference rooms on page 4.

University of Calgary: Campus Map

ST = Science Theatres: Conference lecture rooms
MS = Math Sciences: Team meetings, computer Labs
MSC = MacEwan Student Centre: Food, bookstore, amenities
LRT = Light Rail Transit: Public transit trains
Hotel Alma: Accommodation for participants

Interactive maps: http://ucmapspro.ucalgary.ca/RoomFinder/

Conference Room Guide: Science Theatres

Registration and opening, closing coffee breaks will occur in the foyer, outside ST 141.
Initial group meetings will be in room ST 141.
Team meetings and working coffee breaks are in the adjacent Math Sciences building.
Team meetings will use rooms MS 325, 265, 427, 452, 569. Coffee break in MS 371.
Computer labs are also in the Math Sciences building, MS 317.

Interactive maps: http://ucmapspro.ucalgary.ca/RoomFinder/

General Travel Directions

Here are maps of campus:

Interactive Google map:
http://ucmapspro.ucalgary.ca/RoomFinder/

Full colour non-interactive map:
http://ucmaps.ucalgary.ca/PublicFiles/CurrentMaps/CampusMap_MainCampus_Letter.pdf

Travel Directions from airport:
Calgary is serviced by the Calgary International Airport located 25 minutes from the University of Calgary campus.

Ground transportation between the airport and the University of Calgary campus is available via Uber or taxi:

Associated Cabs: 403-299-1111
Calgary United Cabs: 403-777-1111
Checker Yellow Cabs: 403-299-9999

Calgary Transit (Route No. 300) provides city bus service from the airport to the downtown core, and from there you can take a C-train (light rail transit) to the University. You can also transfer from Route 300 to Routes 20 or 73 for a bus to campus. See Google Maps for details.

Public Transit:
Plan your public transport rides by visiting http://www.calgarytransit.com/ where direction, ticket costs and bus schedules are indicated. The C-train Red line services the university campus.

Parking on Campus:
There are open air lots and parking structures for parking your vehicle, for a fee. Details are online here: https://www.ucalgary.ca/parking/visitorparking

Sunday August 19, 2018

3:00 pm – 11:00 pm	Hotel Check-in for outside participants (Hotel Alma)
7:00 pm – 9:00 pm	Ice Breaker at Jameson's Pub (appetizers and refreshments) 3790 Brentwood Rd NW (North of campus)

Monday August 20, 2018

* Schedule is subject to changes and additions.

8:00 am - 8:45 am	Check-In / Registration & Coffee and Pastries (ST 141 Foyer)
8:45 am - 9:00 am	Welcoming Remarks (ST 141) • **Workshop Organizers** • **Dr. Eric Donovan,** Associate Dean of Science
9:00 am - 9:25 am	**Juris Meija**, National Research Council - Metrology Sample size calculation in homogeneity assessment of certified reference materials
9:30 am - 9:55 am	**Nathan Vadeboncoeur,** Smart Shores Foliage obscured topographical mapping using kinematic GPS survey and aerial drone data
10:00 am - 10:25 am	**Thomas Holloway**, Responsive AI AI for Tactical Asset Allocation
10:30 am - 11:00 am	Coffee break (ST 141 Foyer)
11:00 am - 11:25 am	**Matt McDonald**, Fotech Solutions Inc. Two-dimensional phase unwrapping for distributed acoustic sensing.
11:30 am - 12:00 pm	**Michael Morgan**, GLJ Petroleum Consultants Stochastic modelling of oil and gas economies.
12:00 pm - 12:30 pm	**Introduction** to software tools for workshop Division into teams
12:30 pm - 1:30 pm	**Welcome Lunch** (MacEwan Student Centre) Each participant will each receive a meal card, valid at the campus eateries in the MacEwan Student Centre.

NOTE: We move to the Math Science Building for the team meetings.

1:30 pm - 3:30 pm	Team discussions (Breakout Rooms MS 325, 365, 427, 452, 569)
3:30 pm - 4:00 pm	Coffee break (MS 371)
4:00 pm - 6:00 pm	Team discussions

Tuesday August 21, 2018

NOTE: We move to the Math Science Building for the team meetings.

8:30 am - 10:30 am	Team discussions (Breakout Rooms MS 325, 365, 427, 452, 569)
10:30 am - 11:00 am	Coffee break (MS 371)
11:00 am - 12:30 pm	Team discussions
12:30pm - 1:30 pm	Lunch (MacEwan Student Centre)
1:30 pm - 3:30 pm	Team discussions (Breakout Rooms MS 325, 365, 427, 452, 569)
3:30 pm - 4:00 pm	Coffee break (MS 371)
4:00 pm – 6:00 pm	Team discussions

Wednesday August 22, 2018

8:30 am - 10:30 am	Team discussions (Breakout rooms MS 325, 365, 427, 452, 569)
10:30 am - 11:00 am	Coffee break (MS 371)
11:00 am - 12:30 pm	Team discussions
12:30pm - 1:30 pm	Lunch (MacEwan Student Centre)
1:00pm	Invited Research Talk, by Dr Juris Meija Mathematics in Analytical Chemisry (Math Sciences Room 431)
1:30 pm - 3:30 pm	Team discussions (Breakout Rooms MS 325, 365, 427, 452, 569)

NOTE: We move back to the Science Theatres for the afternoon group presentations.

3:30 pm - 4:00 pm	Coffee break (ST 141 Foyer)
4:00 pm – 6:00 pm	TEAM PROJECT UPDATES (ST 141) 15 minute presentations, by team members - Team 5 - Team 4 - Team 3 - Team 2 - Team 5
6:00 pm - 9:00 pm	Speed Skating Excursion (OPTIONAL) Olympic Oval (University of Calgary main campus) Bring your conference nametag for free entry onto the Oval. Skates and helmets will be available at the Oval.

Thursday August 23, 2018

8:30 am - 10:30 am	Team discussions (Breakout Rooms MS 325, 365, 427, 452, 569)
10:30 am - 11:00 am	Coffee break (MS 371)
11:00 am - 12:30 pm	Team discussions
12:30pm - 1:30 pm	Lunch (MacEwan Student Centre)
1:30 pm – 5:00 pm	Team discussions (Breakout Rooms MS 325, 365, 427, 452, 569)
5:00 pm – 7:00 pm	Barbeque (Outside the Science Theatres, location to be announced)
6:00 --	Working late to wrap up!

Friday August 24, 2018

8:30 am - 9:00 am	Coffee and Pastries (ST 141 Foyer)
9:00 am - 9:25 am	**Project 1 report** (Team 1 - ST 141)
	Sample size calculation in homogeneity assessment of certified reference materials
9:30 am - 9:55 am	**Project 2 report** (Team 2)
	Foliage obscured topographical mapping using kinematic GPS survey and aerial drone data
10:00 am - 10:25 am	**Project 3 report** (Team 3)
	AI for Tactical Asset Allocation
10:30 am - 11:00 am	Coffee break (ST 141 Foyer)
11:00 am - 11:25 am	**Project 4 report** (Team 4)
	Two-dimensional phase unwrapping for distributed acoustic sensing.
11:30 am - 12:00 pm	**Project 5 report** (Team 5)
	Stochastic modelling of oil and gas economies
12:00 pm - 2:00 pm	**Wrap up Lunch** (MacEwan Student Centre)

Thanks to our Partners and Sponsors

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Alberta Government through funding for the Pacific Institute for the Mathematical Sciences, as well as our partner institutes CRM, Fields, and the University of Calgary.

Natural Sciences and Engineering Research Council of Canada
Conseil de recherches en sciences naturelles et en génie du Canada
Canada

Alberta

CRM
FIELDS
PIMS

UNIVERSITY OF CALGARY

Sample size calculation in homogeneity assessment of certified reference materials

Juris Meija (NRC Metrology)

July 1, 2018

1 Background preparation

Statistics, Experimental Design, Optimization and Optimal Design, Data Analysis.

2 Overview: Measurement Standards

National Research Council Canada, Measurement Standards.

Certified reference materials (CRMs) are key to global comparability of measurement results as most laboratory measurements calibrate or verify their measurement results against the international CRMs. CRMs are typically created in a batch of several thousand units and it is inevitable that the property values (such as the mass fraction of arsenic) might vary between the units. This variability is captured as the uncertainty due to homogeneity of the material and is part of the overall combined uncertainty of the certified values. The question of how many samples need to be tested in order to get a reliable estimate of the homogeneity uncertainty is raised here. This question has financial as well as technical implications. If too few samples are analyzed, one might underestimate the homogeneity of the CRM which in turn might lead to diminished trust in future CRMs or a possible recall. Analyzing too many samples, on the other hand, is undesirable as it requires unnecessary expenses, labour, and depletion of the CRM stock.

3 Problem Description

There are two issues to consider.

First, uncertainties are parameters of probability distributions but we do not have a good statistical model on how impurities are typically distributed in materials. Most statistical methods, such as the one-way ANOVA, assume normal distribution of random effects but this cannot be the case here. The concentration levels of various industrial pollutants are often modeled probabilistically using Weibull, lognormal, or gamma distributions. A better understanding of how to model inhomogeneity distributions will, in turn, lead to better understanding of the meaning behind the uncertainties of CRMs.

Homogeneity of a CRM is typically assessed by performing replicate measurements from several CRM units using a balanced nested experimental design. Furthermore, one typically applies random effects statistical model to the data:

$$x_{i,j} = \mu + A_i + e_{i,j} \tag{1}$$

where $x_{i,j}$ represents the jth measurement result for a certain analyte from the ith unit of the CRM, μ is the population mean, A_i is the effect of unit i (due to inhomogeneity of the sample), and $e_{i,j}$ is random variable representing the measurement uncertainty. Typically these two effects (inhomogeneity of the sample and measurement uncertainty) are modeled as normal random variables $A_i \sim \mathrm{N}(0, u^2_{\text{hom}})$ and $e_{i,j} \sim \mathrm{N}(0, u^2_{\text{meas}})$ and the above statistical model is solved for μ, u^2_{hom}, and u^2_{meas} using one-way ANOVA by employing either traditional frequentist methods or by employing Bayesian methods.

We have a database of some 20,000 data containing measurement results of various chemicals in a variety of NRC chemical CRMs. Analysis of this dataset might be helpful when determining which probability distributions might be appropriate to model the underlying problem. In addition, it should be possible to evaluate the homogeneity estimates arising from a variety of statistical models which would approximate the probability density distribution of the analytes using a variety of non-gaussian distributions.

Second, we are seeking optimal parameters for the experimental design of homogeneity studies. International standards on this matter offer only anecdotal guidance. In addition, there are two schools of thought on this question: some argue that the optimal number of units to be analyzed depends on the lot size whereas others say that it does not. For example, the recommended number of units to be analyzed is $max(10, N^{1/3})$ (ISO Guide 35:2017, https://www.iso.org/standard/60281.html) or $min(15, 0.08N)$ (ASTM E826:1986), or simply that one should inspect a number of units between 10 and 30 (USP 905:2011). Strict application of the power analysis to this problem suggests that the sample size is independent from the total lot size. We are seeking to gain more insights on this matter. It is clear that the power analysis plays an important role.

Foliage obscured topographical mapping using kinematic GPS survey and aerial drone data

Nathan Vadeboncoeur (Smart Shores)

July 1, 2018

1 Background preparation

Approximations of functions, signal processing, image processing, statistics, data science.

2 Overview: Environmental mapping and drone data

Environmental data derived from drone-based surveys.

Smart Shores is a science-based tech company focused on making world-class environmental data accessible to everyone, from small not-for-profits to big governments. We provide affordable high-resolution spatial data, ecological analysis and visual media to support responsible and effective stewardship of our natural world.

We help tell the story behind the data by providing both technical (GIS) data and engaging content including interactive 3D environments, detailed maps, short videos, and outreach materials.

A key focus for our many of our clients is environmental management, which is crucial for our quality of life. For example, a well-managed shoreline can support salmon and shorebird populations, provide recreation opportunities and buffer against flooding. However, shoreline development can also have unintended consequences that are harmful to ecosystems, the economy, recreation and culture. We help connect human and ecological needs to support progressive, twenty-first century land management.

3 Problem Description

Smart Shores produces orthomosaics, digital elevation models and 3D point clouds at up to 1 cm resolution and $<$ 5cm precision. We also offer satellite-derived measures of vegetation health and marine plant distributions, and a range of value-added products.

A key technology to achieve this uses a combination of on-demand drone (UAV) and ground survey. This ground-verified drone data provide a cost-effective way to obtain high-precision environmental data. These data support the planning and verification stages of environmental restoration projects, ecological change monitoring, coastal erosion and sediment transport assessment, and evaluation of vegetation health and the presence of invasive plants.

The data we have for this workshop consists of topographical data from both drones and ground surveys. The challenge is that much of the visual drone data is obscured by foliage, which makes the registration with ground-truth a difficult propostion. Some questions we would like to answer in analyzing this data are as follows:

- How many, or how few, ground-based survey points are required to calibrate the foliage-obscured drone data?
- How can the foliage data be adjusted for seasonal variations such as vegetation density or height?
- Are there long-term variations in the data that must be accounted for in calibrating the drone date to ground surveys?

AI for Tactical Asset Allocation

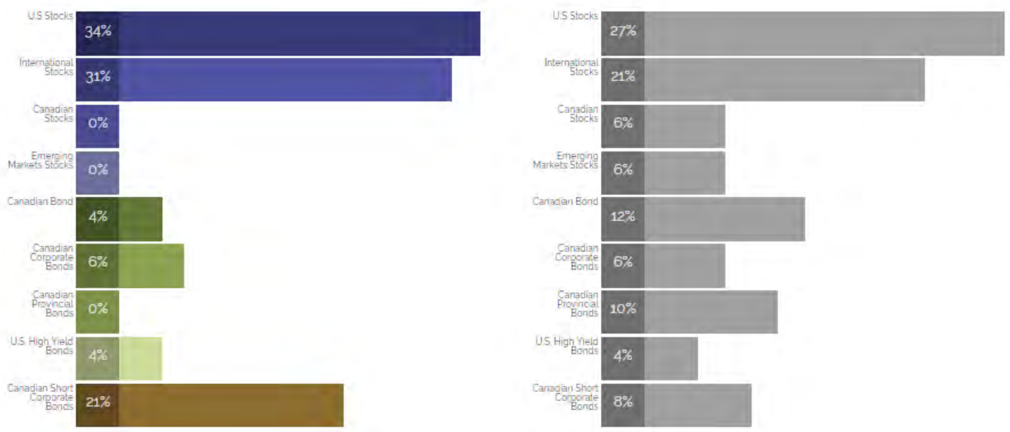
Thomas Holloway (Responsive AI)

August 13, 2018

1 Background preparation

Data Science, Statistics, Statistical Modelling, some Finance.

2 Overview: Wealth Management for Individuals and Families



Can we use data to Responsively adjust asset weights to improve portfolio performance?

Responsive AI is a venture-backed startup in direct-to-client wealth management, hybrid wealth technology (B2B), and data-driven client intelligence research. This summer we are celebrating three years of operation and the close of our US $1.1mm seed financing round.

Responsive's asset management business is differentiated by streamlined user experience, industry-leading cost, and the implementation of data science research for optimized portfolio performance. Eschewing security selection as a source of value, Responsive clients own portfolios of Exchange Traded Funds (ETFs). Each ETF provides exposure to an asset class such as US stocks, foreign stocks, Government bonds, and High Yield Bonds. The Responsive investment process is to reconsider the weights monthly using signals from markets and the economy. Following Samuelson, "The market is micro efficient, but macro inefficient".

3 Problem Description

We will start with a collection of a few hundred time-series data series from the market and economy, coupled with returns in the major asset classes. Using a variety of transformations, such as normalization or computing year-over-year rate of change, the data will be converted to 'factors'. For the asset classes, we are interested in relative returns for asset pairs, for example equity-versus-cash, and within equity, say Canada-versus-US. We will build an infrastructure to test, select, and combine the most reasonable, profitable, consistent, and additive factors. Turnover and portfolio risk are important constraints.

Market factors tend to take the form of valuation, momentum, and mean reversion. Since recessions are such significant divers of relative returns, macroeconomic forecasting is valuable, particularly in combination with market factors. For example, is a given 10% equity market drawdown just a "healthy correction" or the start of true bear market? The answer may be in how real retail sales are holding up.

Markets are adaptive ecosystems of people (and now machines). The modelling challenge is therefore behavioural and biological because the allocation decisions of individuals and investment committees reflexively help determine subsequent returns. Everyone is looking at subsets of the same data, headlines, and prices.

Two dimensional phase unwrapping for distributed acoustic sensors

Matt McDonald

July 1, 2018

1 Preparatory material

Calculus and mathematical modelling, numerical methods, signal processing. Some physics background would be useful.

2 Distributed acoustic sensing

Distributed acoustic sensing (DAS) is a relatively new technology used for measuring strain and vibration along a long cable installed on large linear infrastructures such as railways, roadways, pipelines, underground boreholes, and so forth. The sensor consists of a single-mode fibre optic cable installed along the length of structure under surveillance, and connected to a laser and photodetector. A pulse of light is sent down the fibre, and the reflections scattered back to the detector are analyzed to measure strain on the fibre. Depending on the length of the fibre, the laser may be fired many thousands of times per second, and the Rayleigh backscatter sampled to produce a measurement of the strain at numerous points on the fibre. Knowledge of the refractive index of the fibre allows one to use two-way travel-time to treat the fibre as a series of discrete sensors up to 40 km long, where each position acts somewhat like a two-beam interferometer.

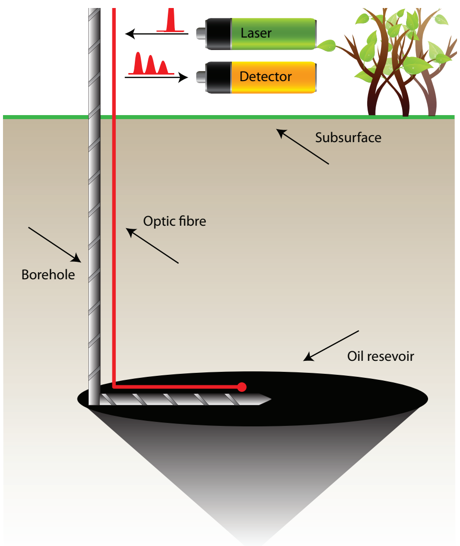


Figure 1: DAS with fibre in borehole.

In terms of typical data volumes collected, a 5 km fibre with a spatial sampling interval of 0.67 meters would translate to roughly 7462 fibre-positions being sampled 20,000 times per second. The signals can be treated as the output of a long string of individual sensors distributed along the cable, with a high frequency sampling rate.

A significant challenge in improving the sensitivity and accuracy of the DAS device is calculating the optical phase of the light signal travelling in the fibre from the measured intensities on the photodetector. Noise and phase wrapping are the main issues that make this difficult.

3 Problem Description

The details of the phase problem arises as follows. In distributed acoustic sensors, a laser pulse is sent along a fibre-optic cable, and the intensity of the backscattered light is measured as a function of time. This backscattered light is related to the elastic strain that the fibre experiences through measurement of the *wrapped* optical phase inside the fibre.

Assume the DAS response can be theoretically modelled as

$$\phi(x,t) = \phi_0(x) + s(x,t) + \eta$$

where $\phi(x,t)$ is the DAS response as a function of space x and time t, $\phi_0(x)$ is random in space, but constant in time, $s(x,t)$ is a continuous function, and η is a random variable. Let $\psi(x,t) = W[\phi](x,t)$ be the *measured* wrapped phase, where $W[\cdot]$ is the wrapping operator $W[\theta] = \theta + 2k\pi$, $k \in \mathbb{Z}$, such that $W[\theta]\ in[-pi, pi)$. The goal is then to recover $s(x,t)$ from $\psi(x,t)$.

The Itoh condition for phase unwrapping in one dimension is that

$$\Delta\phi = W[\Delta\psi]$$

where Δ is the discrete difference operator. Thus, ϕ can be recovered by writing

$$\phi_{i+1} = \phi_i + W[\Delta\psi].$$

Figure 2 shows a sine wave undergoing wrapping and then unwrapping, with and without additive noise. When there is no noise, and the phase is sampled with a sufficiently high sample rate, the phase unwrapping is exact. When the phase changes too much, too quickly, either due to noise or due to too low of a sample rate, the phase unwrapping fails and introduces spurious *jumps* in the data.

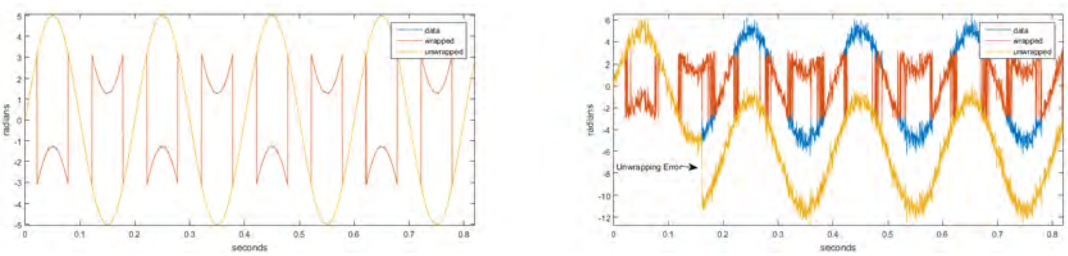


Figure 2: Wrapped and unwrapped phase with and without additive noise.

As distributed acoustic sensors create dense virtual sensors spaced $< 1m$ apart, it is appealing to seek some form of two-dimensional phase unwrapping algorithm that that takes neighbouring sensors into account in order to increase robustness and remove phase unwrapping errors such as those seen in Figure 2. The presence of the random offset of each sensor, represented by the term $\phi_0(x)$, makes this difficult however.

A relatively simple first step in recovering $s(x,t)$ is to estimate the mean of the wrapped phase for each location, subtract that mean-estimate, and then re-wrap the data. This produces

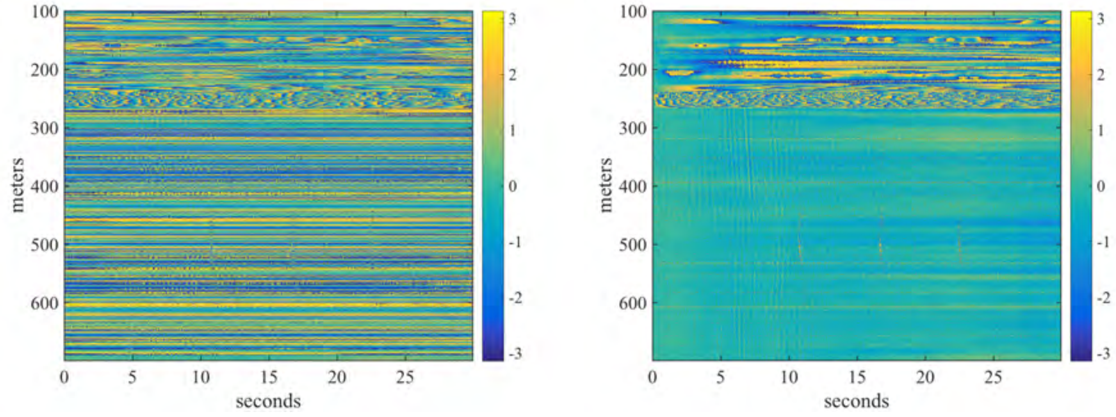


Figure 3: Wrapped 2D phase before and after mean subtraction and re-wrapping.

an estimate of $W[\phi(x,t) - \phi_0(x)] = W[s(x,t) + \eta]$ where small signals less than $\pm\pi$ radians are unwrapped. Figure 3 shows an example of a 2D wavefield recorded before and after subtracting an estimate of the mean of each location computed over 1001 time samples and re-wrapping.

Figure 4 shows a zoomed in region of the data in Figure 3 around a portion of the wavefield that is experiencing wrapping.

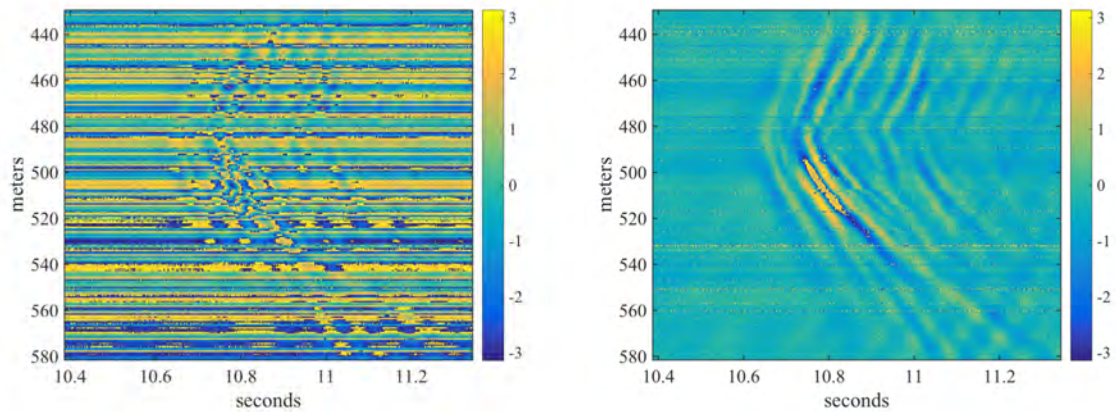


Figure 4: Wrapped 2D phase before and after mean subtraction and re-wrapping.

The general goal of the project is to develop a robust two-dimensional phase unwrapping algorithm, specific to distributed acoustic sensor data. It should be noted that two-dimensional phase unwrapping algorithms already exist for imaging of phase-wrapped surfaces. That is where both dimensions are space in a physical coordinate system. These techniques may prove useful, but the extension to space-time sampling of wavefields could provide further constraint.

Stochastic modelling of oil and gas economics

Michael Morgan (GLJ Petroleum Consultants)

July 1, 2018

1 Background preparation

Differential equations, stochastic modelling, statistical models, time series analysis, econometrics, numerical methods. Some knowledge of financial mathematics and asset pricing would be great.

2 Overview: stochastic economic models

Modelling uncertainty in oil and gas economies.

Many aspects of the oil and gas industry are subject to economic uncertainty: success in finding oil and gas reserves is highly unpredictable, productivity of an individual well can change in seemingly random ways, markets for products can move unexpectedly, future prices and interest rates are unknown, among other challenges. Useful economic models need to account for these uncertainties. Stochastic models provide an approach to these problems by treating certain unknown parameters as random variables and performing an in-depth analysis that takes into account this uncertainty.

Stochastic differential equations (SDEs) are one useful method whose application are of interest to oil and gas economic problems. As most economic analysis is currently conducted using deterministic discounted cash flow models, there is much scope of innovation and experimentation. In particular, there is much scope for improving how uncertainty is mathematically modelled. In these project, we are interested four different general areas:

- Econometrics

- Real Options Analysis
- Geographic Correlation and Temporal Uncertainty
- SDE Solutions to PDEs

3 Problem Descriptions

The four general areas can be approached by specific problems, for which we are looking for specific outcomes. Participants are welcome to try all these problems but are encouraged to choose a subset of these problems and develop more comprehensive solutions complete with specific numeric solutions and supporting code. The use of open source software is preferred, with Python, R, Stan and Octave being particularly well suited to these problems.

3.1 Econometrics

Model construction and estimation of parameters for Canadian energy commodity prices

Most econometric analysis has used price series external to Canada. It is common to consider the price of West Texas Intermediate (WTI) oil or of gas traded at Henry Hub. However, Canadian producers most often receive revenues based on the price of oil at various refineries in Edmonton or on the price of gas at either the AECO trading hub or at Station 2 in British Columbia. A further complication is the relatively large amount of heavy oil produced within Alberta which is sold at a discount to prices of light oil. These local hydrocarbon prices are related to each other and to world benchmarks (such as WTI) but do not necessarily share the same variance or other model structures.

Current practice often models Canadian prices to be simple linear functions of world benchmark prices. However, such models do not allow different mean reversion rates, volatilities or correlations, all of which may be necessary to reflect fundamental physical and economic differences specific to the Canadian market. Exploring price models suitable for Canadian producers would provide improvements in the ability to properly assess price inter-dependencies and uncertainties.

Data for various Canadian price series can be provided by GLJ Publications.

A successful solution would propose an analytical framework suited to the observed dynamics of Canadian energy commodity prices; statistical work supporting the inclusion or exclusion of various factors or terms; calculation of best estimate parameters; forecasts showing equiprobable future outcomes and any computer code required for these calculations.

3.2 Real Options Analysis

Economics of constructing gas processing plants that extract more liquid hydrocarbons

In designing gas processing plants an important factor is the degree to which various higher weight hydrocarbons are separated from methane. This is commonly done by chilling the gas until the higher weight hydrocarbons condense and can be removed as a liquid. As different hydrocarbons have different boiling points, the relative amount of condensation of each hydrocarbon can be controlled by careful choice of temperature. Operators often consider increasing the recovery of higher weight hydrocarbons by constructing gas plants that operate at lower temperatures. This allows the hydrocarbons to be sold as liquids (taking advantage of higher liquids prices) instead of a gas. However, running gas plants at lower temperatures increases operating costs. Moreover, the operating temperature is primarily determined by the design and construction of the plant while the relative value of hydrocarbon liquids to gases is determined by spot and futures markets and is under constant change.

The choice of plant operating temperature can be viewed as a real option. The value of this real option will be determined by whether various hydrocarbon species (ethane, propane, butane, pentane, etc.) are sold as separate liquids or sold within the gas stream, by the incremental cost of constructing plants capable of operating at lower temperatures and by the incremental costs of operating at those lower temperatures. As the prices of various hydrocarbon species are interdependent (and likely co-integrated), a suitable commodity price model will likely be needed.

Constructing a model of the real option of gas plant operating temperature would improve the ability of investors to weigh the relative merits of various plant designs, thereby maximizing the economic return of investments.

Historical data for various hydrocarbon prices series can be provided by GLJ Publications, simplified models for hydrocarbon yields and plant costs can be provided by GLJ Petroleum Consultants.

A successful solution would propose an analytical model that captures the major physical relationships between liquid yield, operating temperatures and operating cost; inclusion of a simplified co-integrated price model; calculation of the optional value of plant construction as a function of input parameters, including temperature and any computer code required for these calculations.

3.3 Geographic Correlation and Temporal Uncertainty

Pipeline supply with discrete correlated declining input sources

Nearly all energy resources are spatially correlated and aggregated using fixed infrastructure. Whether this is the sunlight available for a solar panel, the wind to power a turbine or productivity of a gas well, high output point sources are likely to be located near other high output sources. Within the mining industry such correlations are often characterized using various geostatistical descriptions like variogram models. Such models often use static outputs: the relative presence of minerals for example. However, energy supplies often also have a temporal quality. The wind speed changes throughout the day and gas wells typically produce at higher rates early in their lives. Thus, wise investment decisions on fixed infrastructure will need to consider both spatial and temporal patterns.

It would be beneficial to have a model for pipeline supply that considers a catchment area within which discrete input sources are located (ie gas wells). These input sources are not constant: they decline over time according to uncertain spatially correlated parameters. Such a model should calculate the distribution of number of point sources required, and the pace of adding them, to maintain a constant supply to the pipeline. Such a model should also consider how the distribution changes with increasing certainty, both in terms of the underlying (spatially varying) productivity of the area and in terms of the uncertainty of the decline parameters.

Such a model for pipeline supply would be beneficial when considering development within relatively untested areas, thereby better informing investment decisions and development strategies.

Historical data for various catchment areas can be provided by GLJ Petroleum Consultants.

A successful solution would propose an analytical framework; statistical work supporting the choice of spatial and temporal models; forecasts of various development scenarios, the distribution of point sources required to maintain a constant supply and the impact of certainty of model parameters on this distribution; and any computer code required for these calculations.

3.4 SDE Solutions to PDEs

Stochastic solutions to partial differential equations commonly used to model production from oil and gas wells

There are two main techniques to model production from oil and gas wells: physics-based 3D simulations and approximations based on simple partial differential equations (PDEs). Simple PDE approximations are the more commonly used for economic calculations, especially those with compact analytical solutions. For example, late life wells produced with a constant pressure are often modelled using a simple exponential relationship, $\frac{dq}{dt} = -Dq$ where q is the flow rate and D is a constant rate of decline. The parameters of these models cannot be precisely determined a priori, nor can they be estimated with precision due to the difficulty in precisely controlling producing conditions for accurate flow measurements. As such, forecasts are uncertain.

Our goal here is to model the parameters as stochastic processes and solve the resulting stochastic differential equations using Ito's Lemma. These can then be used to compare with data from actual wells to evaluate how well these stochastic models are working.

Common PDEs used to forecast production include the following:

Arps:

$$\frac{dq}{dt} = -\left(\frac{D_i}{q_1^b}\right) q^{b+1}$$

Power Law Exponential:

$$\frac{dq}{dt} = -\left(D_\infty + D_1 t^{-(1-n)}\right) q$$

Stretched Exponential:

$$\frac{dq}{dt} = -\left(\frac{n}{t}\left(\frac{t}{\tau}\right)^n\right) q$$

Logistic Growth:

$$\frac{dq}{dt} = -\left(\frac{a - an + (1+n)t^n}{t\,(a + t^n)}\right) q$$

Duong:

$$\frac{dq}{dt} = -\left(mt^{-1} - at^{-m}\right) q$$

Stage two of this sub-problem is to define stochastic process to model net cash flow, including royalties. Net cash flow will be revenue minus expenses and taxes. Revenues will be a simple multiple of instantaneous volume produced and instantaneous price. Operating expenses will be a linear function of time plus a simple multiple of instantaneous volume produced and processing cost. Royalties (ie taxes) will be function of average rate and price over an accounting interval, typically a month.

Historical production data and typical economic parameters, including royalty models, can be provided by GLJ Petroleum Consultants.

A successful solution would propose analytical stochastic solutions to a selection of the common PDEs models; calculation of best estimate parameters; forecasts showing equi-probable future production outcomes; forecasts showing the distribution of expected future revenues and any computer code required for these calculations.