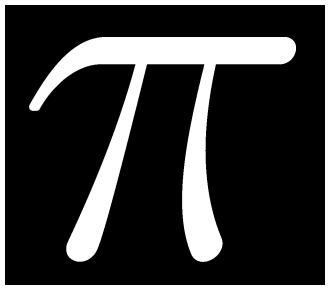


Pacific Institute for the Mathematical Sciences

http://www.pims.math.ca
pims@pims.math.ca

Proceedings of the second

PIMS Graduate Industrial Math Modelling Camp

Cosponsored by:

The Natural Science and Engineering Research Council of Canada

and

The British Columbia Information, Science and Technology Agency

FOREWORD BY THE PIMS DIRECTOR

The PIMS **Graduate Industrial Math Modeling Camp (GIMMC)** is held every year in one of the PIMS universities as part of the annual PIMS Industrial Forum. It is part of PIMS commitment to providing training for young mathematical scientists who are either pursing careers in academia or in industry.

The goal of the GIMMC is to provide experience in the use of mathematical modelling as a problem solving tool for graduate students in mathematics, applied mathematics, statistics and computer science. In addition to this it helps prepare them for the **Industrial Problem Solving Workshop** which is the other component of the PIMS Industrial Forum.

At the workshop students work together in teams, under the supervision of invited mentors. Each mentor poses a problem arising from an industrial or engineering application and guides his or her team of graduate students through a modelling phase to a resolution.

The 1999 GIMMC, which was the second, was held at University of Alberta, May 24–28, 1999. Twenty-seven students came from eight universities across Canada to work with six mentors from industry. These mentors came from Alberta Energy Utilities Board, Telecommunications Research Laboratories, Syncrude, Alberta Research Council, Enbridge Pipelines Ltd and Lockheed Martin Canada.

It my pleasure to announce that the workshop was an unqualified success for all the six project teams.

I want to extend my thanks to the organisers (Abel Cadenillasi, Douglas Kelker, Mike Kouritzin, Henry Leung, Bryant Moodie, Bruce Sutherland, Yanhong Wu), mentors (Stefan Bachu, Wayne Grover, Mike Lipsett, John Oliver, Don Scott, Pierre Valin) and students involved in this workshop. Future camps are something I truly look forward to.

Dr. Nassif Ghoussoub, Director
Pacific Institute for the Mathematical Sciences

PREFACE

This year's GIMMC adopted a novel approach to introducing graduate students to the mathematical techniques that arise in industrial R&D. Rather than tutoring students through solutions to artificially posed problems or to problems that have already been solved, this camp asked students to tackle a set of six, unsolved, modelling problems brought to them by industrial mentors. These problems arose directly from the R&D activities of the mentor's organisation. Based on team-project approach, each group of students, a faculty partner and an industrial mentor, worked together to solve one of the six problems.

Judging from the written feedback, the Second PIMS Graduate Industrial Math Modelling Camp was an unqualified success for all the six project teams. Everyone praised the camp's impeccable organisation. The Industrial Mentors were unanimous in their praise of the camp's concept and execution. The industrial mentor from Lockheed Martin Canada wrote that the results of the camp exceeded his expectations. The industrial mentor from Telecommunications Research Laboratories, one of Canada's leading industrial telecommunications research organisations, was unequivocal in his assessment: "I think it's all a terrific idea. I overwhelmingly salute the whole initiative". The graduate students got the satisfaction of making meaningful progress on mathematical questions of timely interest to industry. As one student from Montreal put it: "I really enjoyed working on a concrete problem". Mathematicians tend to work alone. Doing research as part of team was a new and rewarding experience for the students. For one of the students in Project 4, "It was amazing how well we worked together ... we got a lot out of it and I know we are all happy with our report". Another student in Project 1 appreciated the experience of "learning how to contribute ideas with other students and write a report together". Through the interaction with the graduate students and faculty partners, the industrial mentors developed new perspectives on their modelling problems. But most important, the camp exposed all the participants to the rewarding possibilities that collaboration between mathematicians and industry can bring.

As with any undertaking, success starts with the quality of the people involved. GIMMC was fortunate to have had enlisted senior members from some of Canada's most prominent organisations to act as "industrial mentors".

The faculty partners were critical to the success of this camp. They provided the needed bridge between the mentor's industrial expertise and the graduate students' mathematical skills. PIMS wants to give special thanks to all the faculty partners for their selfless participation.

Finally, the camp's success was assured by the participation of bright and motivated graduate students. Twenty-seven students were selected for this year's camp. They were a diverse group whose backgrounds ranged from pure and applied mathematics to statistics. They were also geographically heterogeneous group: four from Quebec, one from Ontario, one from Saskatchewan, eight from Alberta, and thirteen from British Columbia. PIMS is particularly pleased that a camp on Industrial Mathematics attracted as many women as it did men.

Contents

Chapter 1

Regional - Scale Variation of Rock Properties in the Viking Formation in Alberta

Participants: Stefan Bachu (Mentor), Andrea Amariel, Brenda Hawkins, Emmanuel Ngimbi Ngembo, Shelly Pinder, Grace So.

PROBLEM STATEMENT: The Cretaceous sedimentary succession in the Alberta basin is comprised of sandstone and shale formations. The main sandstone units are, in descending order from the surface: Paskapoo, Scollard, Belly River, Milk River, Dunvegan, Cardium, Viking and Mannville. Oil and gas exploration and production has produced a wealth of information regarding porosity and permeability in these units. The data are filed in electronic form with the Alberta Energy and Utilities Board and are publicly available. So far no analysis of rock properties in these units has been performed at the basin scale. The main issue is to identify if there are regional-scale trends in porosity and permeability variations that would allow prediction of these rock properties for areas and strata lacking data. The Viking Formation has been selected for testing for trends in rock properties and for the development of methodology to be applied in the analysis of the other sandstone units in the Alberta basin.

More specifically, the objective is to statistically analyze permeability and porosity distributions in the Viking Formation in order to identify at the regional scale any of the following:

- if there is any correlation of rock properties with depth;
- if there is any correlation between the two rock properties;
- if there is a particular spatial distribution of rock properties.

1.1 Introduction

Except for the northeast corner of Alberta where Precambrian crystalline rocks are exposed at the surface, Alberta's subsurface is comprised of a wedge of sedimentary rocks that thickens from a zero-edge in the northeast to more than 6000 m in the southwest at the Rocky Mountain foothills. One of the main characteristics of sedimentary rocks is porosity, ϕ, which is defined as the ratio of "empty" space in a rock volume to the total rock volume. This empty space is actually filled, or saturated, with fluids, mostly water. In hydrocarbon reservoirs, abundantly found in Alberta, the pore space is filled with oil or gas. While essential for the existence of fluids in the sedimentary rocks, porosity is not indicative of the ability of fluids to flow through the pore space. Permeability, usually denoted by k, is the rock property that indicates the ability of fluids to flow through rocks when a hydraulic gradient is applied as a result of either natural processes, such as elevation differences, or of human activity, such as pumping or injection of fluids. Permeability is an essential rock property for pumping oil and gas out of hydrocarbon reservoirs, and for injecting residual water, liquid wastes and carbon dioxide into deep saline aquifers and depleted reservoirs. Both rock properties are geo-spatially distributed, exhibiting a dependence on location and depth.

As siliciclastic sediments (sand, mud, clay) are deposited in a sedimentary basin, they are mechanically compacted by the weight of the overburden, such that their porosity decreases with depth from about 45% at deposition to a few % at great depths. The following empirical relation, Athy's Law, was observationally found to be valid in many cases:

$$\phi = \phi_0 e^{-bz}, \tag{1.1}$$

where z is depth, ϕ_0 is the sediment porosity at deposition ($\sim$ 45%), and b is an empirical constant.

Permeability is linked to porosity, but not in a simple direct way. Sediments with high porosity such as clays and shales, have extremely low permeability, making them aquitards (rock formations unable to transmit water on a human time scale). Other rocks such as sandstones may have a smaller porosity, but a significantly higher permeability, making them aquifers (if they contain water) or reservoirs (if they contain oil or gas) from which fluid can be extracted. For the same lithological unit, e.g. sandstones, permeability may be highly variable, by several orders of magnitude, from 10^{-15} to $10^{-10} m^2$. The oil industry uses for permeability another unit, the Darcy (1 D = $10^{-12} m^2$). Permeability displays a much weaker correlation with depth, but in some cases an empirical relation with porosity was found, of the form [7]:

$$log(k) = A\phi + B, \tag{1.2}$$

where k is expressed in Darcies, and the coefficients A and B have values around 15 and -3, respectively.

The values of porosity and permeability must be determined through experiments and field measurements. Porosity is measured in the laboratory on plug samples, several cm long, extracted from core. Unlike variables like temperature and pressure, permeability cannot be directly measured, and it has to be back-calculated from the measurement of other variables, such as pressure and flow rate, and application of Darcy's Law for flow in porous media. Permeability can be determined in the laboratory on the same plug samples as porosity. Also, it can be determined in the field in either pump or drillstem tests. In the case of field tests, the volume of rock sampled is several orders of magnitude greater than in the case of laboratory determinations on core plugs. Porosity and permeability values measured at the plug scale (10^{-2} m) can be scaled up to the well, or drillstem test, scale ($10^0 - 10^1$ m) using effective-averaging procedures [3, 5]. Because of sampling bias, measurement scale and rock heterogeneity, the permeability values measured by the two methods are different.

Porosity and permeability are known only at specific locations. Natural porous formations in sedimentary basins are heterogeneous, i.e. they display porosity and permeability variability at various scales. This irregular and complex variability defies a precise quantitative description because of: 1) insufficiency of information, and 2) lack of interest in knowing the structure and flow field in every minute detail, which, even if known, cannot be handled. Prediction of porosity and permeability is important for the exploration and production of oil and gas, and for the selection of sites for injection of residual water and liquid wastes.

	X	Y	THETA	DIST	GROUND	CORDEPTH	ELEV	SCALEPER	POROS	CORLEN
X	1.000	-.507	-.772	.768	-.469	-.690	.727	.147	.636	-.047
Y		1.000	.935	.136	-.294	-.068	.000	.190	.070	-.082
THETA			1.000	-.200	-.045	.217	-.284	.069	-.197	-.035
DIST				1.000	-.802	-.903	.897	.312	.828	-.075
GROUND					1.000	.893	-.826	-.288	-.698	.064
CORDEPTH						1.000	-.991	-.278	-.805	.029
ELEV							1.000	.263	.805	-.018
SCALEPER								1.000	.382	-.026
POROS									1.000	-.026
CORLEN										1.000

Table 1.1: Pearson Correlations among variables from core analysis data.

1.2 Method

Alberta Geological Survey provided three data files, *core.dat*, *dst.dat*, and *dstcore.dat*, which were read into the statistical application SPSS 8.0. Inconsistencies in the data were resolved, and then multiple linear regression strategies and statistical inferencing techniques were used to seek correlations among the variables, to attempt to verify relations 1.1 and 1.2, to produce useful models, and to address the specific questions posed by Alberta Geological Survey.

The geographical location of a well site is specified in the data as X and Y coordinates (in centimeters, transformed from the latititude - longitude system of coordinates to a UTM map at scale 1:2,500,000) from the origin at 49°N and 126°W. This suggested the introduction of two other variables, the polar coordinates of the well location, calculated as the distance from the origin, labelled DIST, and the angle from the polar axis, labelled THETA.

1.3 Porosity and Permeability Measured in Core

The first file, *core.dat*, has 4079 records of plug sample data, with measurements of permeability and porosity already scaled up from the plug to the well scale. We removed 133 records which were found to be from other formations, leaving 3946 valid records. Another 196 records for which the permeability measurement was zero, were removed for any work with that variable, so that those results are based on 3750 records. The variables used for investigation are:

X and Y	geographical coordinates in cm
DIST	distance = $\sqrt{X^2 + Y^2}$
THETA	$\tan^{-1}(Y/X)$
GROUND	ground elevation (m)
CORDEPTH	depth (m) of the cored and analyzed interval
ELEV	elevation (m) above sea level of the cored and analyzed interval
SCALEPER	effective well-scale permeability in millidarcies (1md = 10^{-15} m^2), scaled up from plug-scale values
GEOAVE	geometric average of plug-scale permeability values, same units
POROS	effective well-scale porosity, scaled up from plug-scale values
CORLEN	length (m) of the cored and tested interval

Pairwise correlations were produced and examined for links between variables (Table 1.1). Several of the significant spatial correlations are not surprising: GROUND = ELEV + CORDEPTH, so of course they are related. We also expect DIST to be correlated with ELEV, CORDEPTH, and GROUND because the elevation contour lines on the maps form, approximately, radial arcs to the coordinate origin. A scatterplot of ELEV vs DIST in Figure 1.1 depicts the strength of the relationship between these two variables. We also note that POROS is strongly correlated with the spatial variables, whereas neither of the permeability variables, SCALEPER nor GEOAVE, are significantly related to the spatial variables. There is only a weak correlation of POROS with the permeability variables.

The large negative correlation of DIST with GROUND is due to the topography decreasing northest-

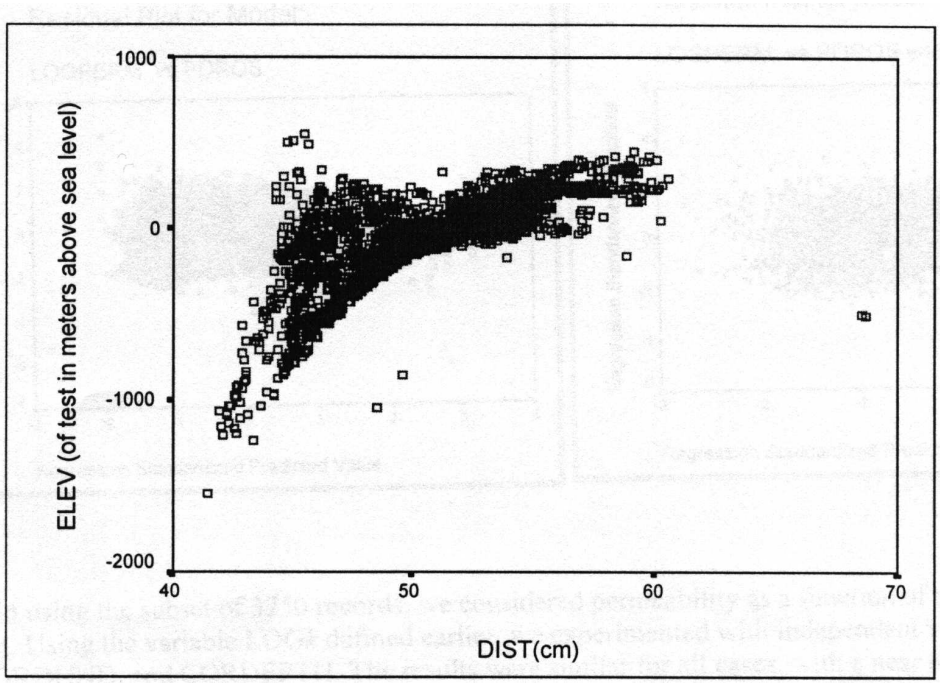


Figure 1.1: Elevation of Tested Interval vs Distance from Origin.

ward (a geographic-erosional feature). The large negative correlation between DIST and ELEV is due to the elevation of the formation increasing northeastward (a depositional feature). Since porosity decreases with depth, the large correlations of POROS with DIST and CORDEPTH is closely related to the relations of DIST with GROUND and DIST with ELEV.

First, we looked at permeability as a function of porosity. Using the subset of 3750 records, we defined a new variable, LOGk, as the logarithm base 10 of (SCALEPER $\div$ 1000) and tried to relate it to POROS, as in Relation 1.2 (This relation was suggested by Rubey and Hubbert [7]):

$$\text{LOG}k = -3.398 + 11.126 * \text{POROS}, \quad \text{with } \text{R}^2 = 0.450.$$

Comparing this equation to Relation 1.2, we observe the constant term, -3.398, is very close to the value of -3 reported by Bethke and Marshak [2], and the coefficient of POROS, 11.126, is not far off their observed value of 15.

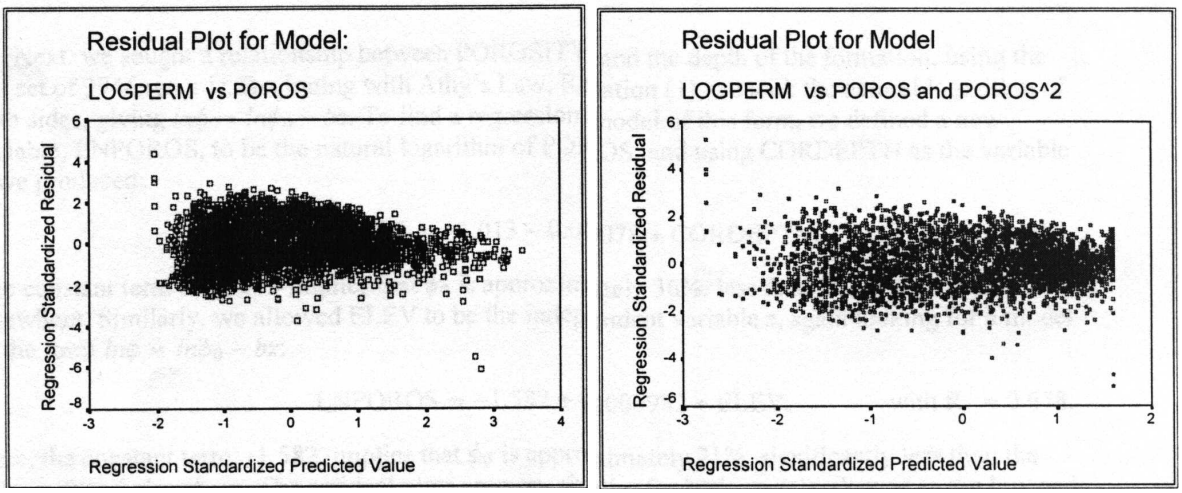
The standardized residuals for this model are not evenly distributed, as seen in Figure 1.2(a). A curve in the scatterplot of LOGk vs POROS suggested the inclusion of POROS^2 in the model. With this improvement we obtain a slightly better R^2 value and the scatterplot of the standardized residuals versus the standardized predicted values displays a more even distribution (Figure 1.2(b)).

$$\text{LOG}k = -3.995 + 20.559 * \text{POROS} - 29.461 * \text{POROS}^2, \quad \text{with } \text{R}^2 = 0.474.$$

Again using the subset of 3750 records, we considered permeability as a function of depth or elevation. Using the variable LOGk defined earlier, we experimented with independent variables ELEV, GROUND, and CORDEPTH. The results were similar for all cases, with a near normal distribution of residuals (Figures 1.3(a) and 1.3(b)), but even the best model has a very low R^2 value.

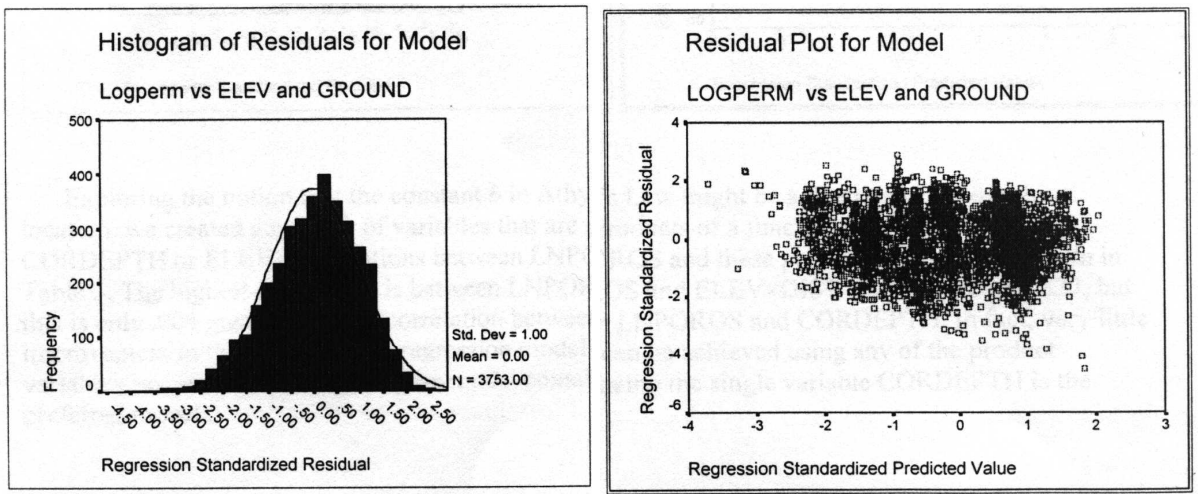
$$\text{LOG}k = -0.597 + 0.00122 * \text{ELEV} - 0.00068 * \text{GROUND}, \quad \text{with } \text{R}^2 = 0.322.$$

Next, we sought a relationship between POROSITY and the depth of the formation, using the full set of 3946 records. Beginning with Athy's Law, Relation 1.1, we took the natural logarithm of both sides,



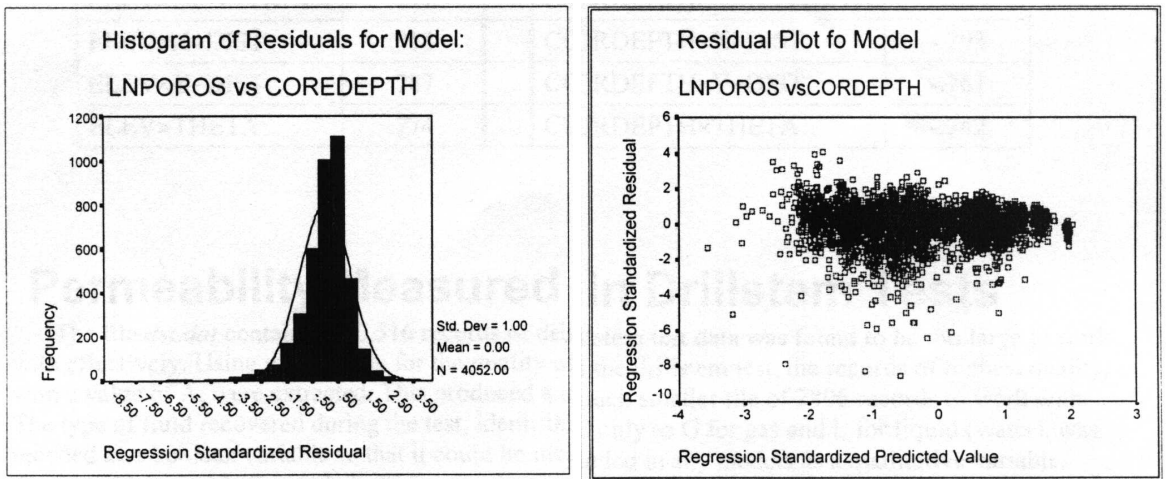
(a) (b)

Figure 1.2: Standardized Residual Plots



(a) (b)

Figure 1.3: Standardized Residual Plots



(a) (b)

Figure 1.4: Standardized Residual Plots

giving $ln\phi = ln\phi_0 - bz$. To find a regression model of this form, we defined a new variable, LNPOROS, to be the natural logarithm of POROS, and using CORDEPTH as the variable z, we produced:

$$\text{LNPOROS} = -1.013 - 0.00078 * \text{CORDEPTH}, \quad \text{with } R^2 = 0.678.$$

The constant term of -1.013 implies that ϕ_0 is approximately 36%, less than the 45% as found elsewhere. Similarly, we allowed ELEV to be the independent variable z, again looking for a model of the form $ln\phi = ln\phi_0 - bz$:

$$\text{LNPOROS} = -1.582 + 0.000982 * \text{ELEV}, \quad \text{with } R^2 = 0.678.$$

Here, the constant term, -1.582, implies that ϕ_0 is approximately 21%, significantly less than the 45% as found elsewhere. The residual plots are very similar for both models, skewed to the low end of values, and revealing a departure from normality (Figures 1.4(a), 1.4(b)). Based on the knowledge of the proosity of sands at deposition, the first model is better than the second one.

Exploring the notion that the constant b in Athy's Law might be a function of geographical location, we created a number of variables that are products of a function of X and Y, with CORDEPTH or ELEV. Correlations between LNPOROS and these product variables are shown in Table 1.3. The highest correlation is between LNPOROS and ELEV×DIST with a value of 0.827, but this is only .004 greater than the correlation between LNPOROS and CORDEPTH. In fact, very little improvement in the R^2 value of a regression model can be achieved using any of the product variables, so, as mentioned above, the model containing the single variable CORDEPTH is the preferred model.

1.4 Permeability Measured in Drillstem Tests

The file *dst.dat* containing 10,516 records of drillstem test data was found to be too large to work with effectively. Using the variable for the quality of the drillstem test, the records of highest quality, with a value of A, were extracted. This produced a much smaller file of 2806 records to work with. The type of fluid recovered during the test, identified only as G for gas and L for liquid (water), was recoded as a boolean variable so that it could be included in any models as a qualitative variable. The variables used in the statistical investigations are:

	LNPOROS			LNPOROS
ELEV	.824		CORDEPTH	-.823
ELEV×X	.819		CORDEPTH×X	-.781
ELEV×Y	.792		CORDEPTH×Y	-.751
ELEV×X×Y	.817		CORDEPTH×X×Y	-.780
ELEV×(X+Y)	.827		CORDEPTH×(X+Y)	-.822
ELEV×(X-Y)	.641		CORDEPTH×(X-Y)	-.362
ELEV×DIST	.827		CORDEPTH×DIST	-.821
ELEV×DIST^2	.827		CORDEPTH×DIST^2	-.811
ELEV×X÷Y	.757		CORDEPTH×X÷Y	-.583
ELEV×Y÷X	.732		CORDEPTH×Y÷X	-.681
ELEV÷DIST	.817		CORDEPTH÷DIST	-.822
ELEV×DIST×THETA	.779		CORDEPTH×DIST×THETA	-.730
ELEV×DIST÷THETA	.777		CORDEPTH×DIST÷THETA	-.582
ELEV×X÷DIST	.816		CORDEPTH×X÷DIST	-.793
ELEV×Y÷DIST	.787		CORDEPTH×Y÷DIST	-.761
ELEV×THETA	.774		CORDEPTH×THETA	-.742

Table 1.2: Correlation between LNPOROS and product variables from *core.dat*

	X	Y	DIST	GROUND	DSTDEPTH	ELEV	PERM	GAS	DSTLEN
X	1.000	-.409	.646	-.296	-.533	.566	.052	.145	.006
Y		1.000	.404	-.551	-.290	.143	.081	.141	.097
DIST			1.000	-.743	-.813	.740	.115	.215	.049
GROUND				1.000	.784	-.593	-.129	-.204	-.076
DSTDEPTH					1.000	-.965	-.095	-.093	-.012
ELEV						1.000	.069	.034	-.017
PERM							1.000	.488	-.177
GAS								1.000	.096
DSTLEN									1.000

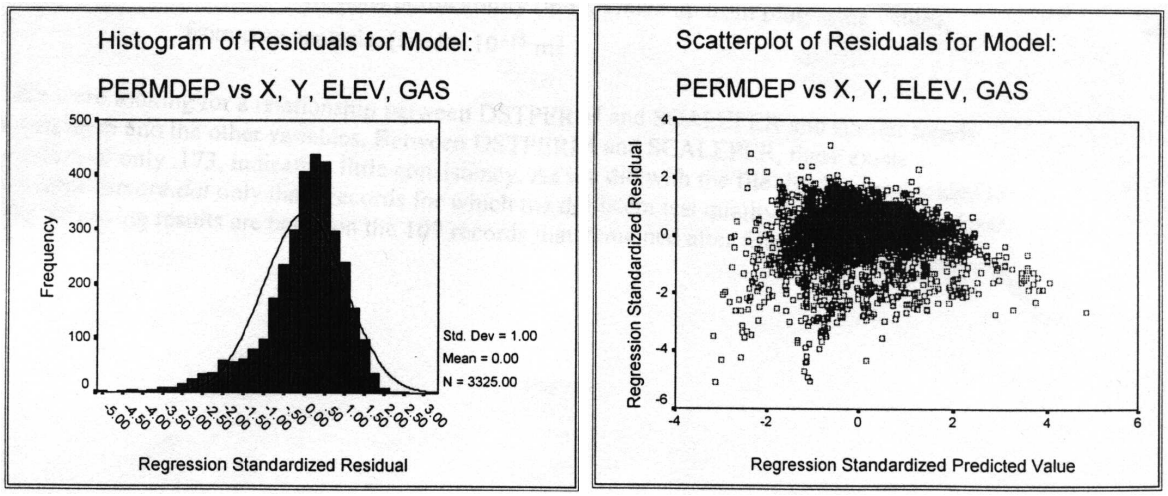
Table 1.3: Pearson Correlations among variables from drillstem test data.

X and Y	geographical coordinates (still in centimeters)
DIST	distance = $\sqrt{X^2 + Y^2}$
GROUND	ground elevation (m)
DSTDEPTH	depth (m) of the drillstem tested interval
ELEV	elevation (m) above sea level of the tested interval
PERM	logarithm base 10 of permeability (m^2)
GAS	boolean values: 1 for gas, 0 for liquid
DSTLEN	length (m) of the tested interval

As with the previous file, bivariate correlations were produced and examined. In Table 1.3 we again see the expected strength of correlation between the GROUND, CORDEPTH, ELEV, and DIST variables. We note that PERM is not strongly correlated to any of the spatial variables, but that it does show a correlation of 0.488 with GAS.

To further explore the significance of the type of fluid present in the sandstone, we divided the data set into two groups, by fluid type. The result was a set of 733 records for which GAS = 0, or the fluid is a liquid, and another set of 2,073 records for which GAS = 1. Using a t-test to compare the mean PERM of the two samples, we found the difference between the sample means to be 1.70, with very strong statistical evidence that the mean PERM is different for GAS = 0 and GAS = 1 (the t test statistic is 36.69).

Multiple linear regression was attempted using both stepwise and backward elimination approaches to identify the important predictors of permeability. In early attempts the residuals always displayed increasing variability. To counter this result, we attempted several transformations of the permeability variable, the most fruitful of which was PERMDEP, defined as PERM divided by the square root of DST-DEPTH. We also experimented with several interactive terms, but found only one useful: GAS×ELEV,



(a) (b)

Figure 1.5: Standardized Residual Plots

which is equal to ELEV when GAS=1, otherwise 0. Using PERMDEP as the dependent variable, we found that both X and Y were entered and retained as independent variables in the models, but DIST was always excluded or removed. We present here the three best models.

$$\text{PERMDEP} = -0.368 - 0.0024 * \text{Y} + .082 * \text{GAS} - 0.0021 * \text{DSTDEPTH} \\ -0.00068 * \text{DSTLEN} - 0.00049 * \text{DIST},$$

with $R^2 = .698$. This model was improved somewhat by the addition of the interactive term GAS×ELEV:

$$\text{PERMDEP} = -0.393 - 0.00233 * \text{Y} + 0.0826 * \text{GAS} + 0.00025 * \text{DSTDEPTH} - 0.00068 * \text{DSTLEN} \\ -.00518 * \text{DIST} + .00009 * \text{GAS} \times \text{ELEV}$$

with $R^2 = .706$. Inclusion of the variable GROUND improved the R^2 value again:

$$\text{PERMDEP} = -0.284 - 0.000293 * \text{Y} + 0.00028 * \text{GAS} + 0.081 * \text{DSTDEPTH} + 0.00071 * \text{DSTLEN} \\ -0.00567 * \text{DIST}, - .0001 * \text{GAS} \times \text{ELEV}$$

with $R^2 = .710$. Unfortunately, each of these three models produces a large number of negative standardized residuals below -3, that only increases with each additional independent variable. Residual plots for all three models appear very similar so only those for the first model are shown here, as Figures 1.5(a) and 1.5(b). These plots display the skewed distribution of the residuals.

1.5 Permeability Measured in Drillstem Tests and Core

The third file, *dstcore.dat* contains 441 records, representing the 441 locations for which we have both the core analysis and drillstem test data, or, in other words, it is the intersection of the other two files, by location. Each record then supplies measurements of all the variables defined for *dst.dat* as well as those for *core.dat*. Some of the variables, such as ground elevation, are redundant in this file, and were disregarded.

	DSTPERM	SCALEPER
X	-.043	.103
Y	.328	.243
DIST	.221	.318
GROUND	-.286	-.252
DSTDEPTH	-.202	-.234
ELEV	.181	.259
GAS	.014	-.008
POROS	.291	.043
DSTPERM	1.000	.356
SCALEPER	.356	1.000
GEOAVE	.285	.952

Table 1.4: Correlations from *dstcore.dat*

	DSTPERM	SCALEPERM
Minimum	0.3	0.1
Maximum	4172.5	6240
Interquartile Range	30	124.9
Median	58.5	4.6
Mean	180.6	57.4
Standard Deviation	257.2	468.4

Table 1.5: Descriptive Statistics

Alberta Geological Survey constructed this file to test the two measurements of permeability for consistency. A core analysis, scaled up from plug-scale values, and a drillstem test were converted to permeability in millidarcies. We considered the following set of variables in our analysis:

X and Y	geographical coordinates
DIST	distance = $\sqrt{X^2 + Y^2}$
GROUND	ground elevation (m)
DSTDEPTH	depth (m) of the drillstem tested interval
ELEV	elevation (m) above sea level of the tested interval
GAS	boolean values: 1 for gas, 0 for liquid
DSTLEN	length (m) of the tested drillstem interval
CORLEN	length (m) of the cored and tested interval
POROS	effective well-scale porosity, scaled up from plug-scale values, from core analysis
DSTPERM	logarithm of permeability (m^2), from drillstem test
SCALEPER	effective well-scale permeability (md), scaled up from plug-scale values, from core analysis, (1md = 10^{-15} m^2)

We were looking for a relationship between DSTPERM and SCALEPER and similar trends between them and the other variables. Between DSTPERM and SCALEPER, there exists a correlation of only .173, indicating little consistency. As we did with the file *dst.dat*, we decided to select from *dstcore.dat* only those records for which the drillstem test quality is of the highest level, A. The following results are based on the 109 records that remained after filtering.

Bivariate correlations are shown in Table 1.4. There are some deviations in the trends between the permeability variables and the other variables. For the variables X and GAS, the correlations with DSTPERM and SCALEPER even have different signs. The correlation of DSTPERM with SCALEPERM is .356, better than that for the whole data set, but still weak. The descriptive statistics in Table 1.4 show significant differences between the two variables, which are supposed to be measuring the same property. The scatterplot of values in Figure 1.6 illustrates this disparity.

These results indicate that there is not a consistent relation between permeability measurements in the drillstem test and the core analysis. Inference based on a combination of the two data sets is

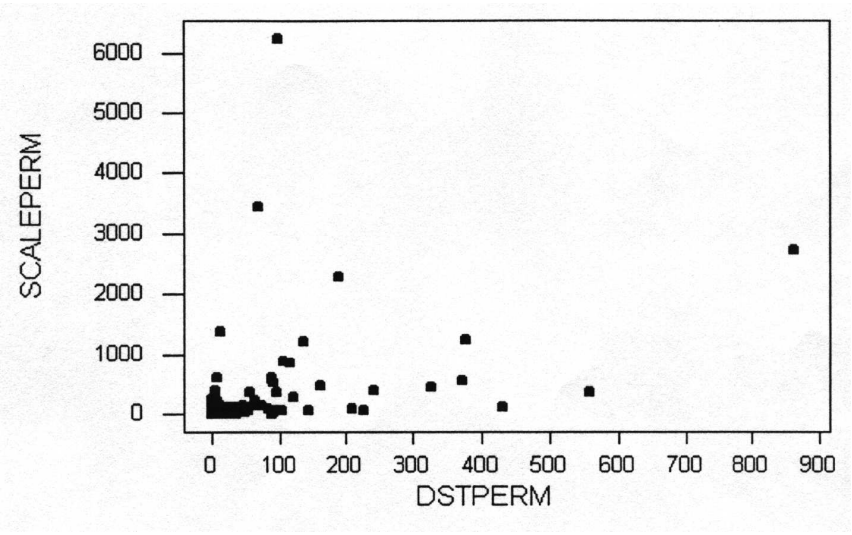


Figure 1.6: Scatterplot of the two methods of measuring permeability.

inappropriate and *dst.dat* and *core.dat* should be analysed separately.

1.6 Conclusion

The two measurements of permeability, the drillstem test at the well scale and the scaled up measurement from the core analysis, are found to be inconsistent. For this reason, the two files were analysed separately, with no mixing of the data sets.

The regression models developed for permeability measured in drillstem tests were less than satisfactory because of the non-normal distribution of the residuals. The presence of gas in the sandstone appears to be a more significant factor of permeability than geographic location or depth.

Permeability measured in core and scaled up is modelled by Relation 1.2, although with an R^2 value of only 0.442. These data do not confirm the coefficients previously observed for Relation 1.2.

Regression reveals a much stronger relationship between porosity and elevation above sea level or, almost equivalently, depth below ground level, as modeled by Relation 1.1. Approximately 70% of the increase in porosity is explained by an increase in elevation or a decrease in depth. These data, however do not support or contradict the value of 45% as the sediment porosity at deposition.

Our investigations fail to indicate specific locations for CO_2 sequestration.

Permeability measured in the core (10^{-2} metres scale) is not representative for formation permeability, as measured in drillstem tests ($10^0 - 10^1$ metres), therefore not for the formation ($10^2 - 10^4$ metres).

Chapter 2

Performance Modelling of a Novel Scheme for Telecommunications Network Synchronization

Participants: Wayne Grover (Mentor), Tamar Dakic, Scott MacLachlan, Todd Oliynyk, Paul Wiebe.

PROBLEM STATEMENT:

Chapter 3

Oilsand Dynamics

Participants: Mike Lipsett (Mentor), Vincent Lemaire, Devom McCrea, Luz Palacios, Peilin Shi.

PROBLEM STATEMENT: In surface mining operations, large face shovels are used to break out ore and load trucks which transport the material to a mill for processing. In the oilsands industry, the processing results in synthetic crude oil. At Syncrude, over 150 million tons of oilsand is excavated each year to produce 80 million barrels of oil (about 12% of Canada's energy supply). Efficient excavation operations are thus very important for this economy of scale. Several factors are related to efficient excavation, including shovel/bucket design, shovel-soil interactions, equipment design, and automated operations.

Surprisingly, this is not a solved problem. Despite millions of earthmoving machines that carve the earth, we don't really understand the physics of moving dirt. We here examine the specific problem of bucket-soil interaction.

3.1 Introduction

In the oilsands, the performance of a shovel is often much poorer than its equivalent performance digging blasted hard rock or coal. Oilsand is a mixture of sand, clay, water, and bitumen. It has the consistency of old asphalt. It is extremely abrasive. In the summer it oozes tar, and sticks to anything it touches. In the winter, it freezes solid and becomes hard to break. It has very high friction against steel.

The difficulty in modeling tool interaction is not limited to the oilsands industry. There exist no good models for how bulk solids break and flow into a bucket or hopper. Efforts to date have concentrated on empirical formulae for capturing the important physical parameters of a machine in simple (and thus inexact) terms. Empirical formulae apply to specific machine types, and neither scale well for new larger machines of the same type nor translate well to other machine types.

3.2 Statement of Problem

We examine oilsand excavation using a front end bucket to extract material from a sloped wall of oilsand. In practice, it has been observed that the timing and motion of the scooping action greatly affects the amount of material recovered in a single scoop. We wish to model this extraction process and examine how various parameters affect the volume of material recovered. Parameters of interest include the soil properties, the angle of the wall of oilsand relative to vertical, the angle of bucket insertion, and the forces applied to the bucket.

3.3 Method of Approach

To model the bank of oilsand, we make use of a specialized program called *FLAC* (**F**ast **L**agrangian **A**nalysis of **C**ontinua). *FLAC* is a two dimensional explicit finite difference program for engineering mechanics computation. It simulates the behaviour of structures built of soil, rock or other materials that may undergo plastic flow when their yield limits are reached. We began by constructing a quadrilateral to model the oilsand, which looks like the graph in Figure 3.1.

Model geometry was specified, as well as soil properties. The following values were used:

parameter	description	typical value
lengthb	depth of insertion of shovel	1m
theta	angle of sand face to vertical	18^o
phi	angle to normal of inserted shovel	10^o
b	bulk modulus	200 MPa
s	shear modulus	200 MPa
c	cohesion	1MPa
t	tension	5.67 MPa
f	friction	10^o
di	dilation	10^o
d	density	2000 kg/m^3

We assumed that the shovel was already inserted into the sand.

Force was applied to the inserted part of the bucket with a linear variation to simulate the scooping motion, as shown in Figure 3.2.

The graphs in Figure 3.3 are obtained after 100 (upper left), 200 (upper right), 500 (lower left) and 1000 (lower right) steps:

The graph of the change of the vertical displacement after 500 time steps is shown in Figure 3.4.

A difference can be seen when changing theta to 15^o. The graph of the change of the vertical displacement after 500 time steps is shown in Figure 3.5.

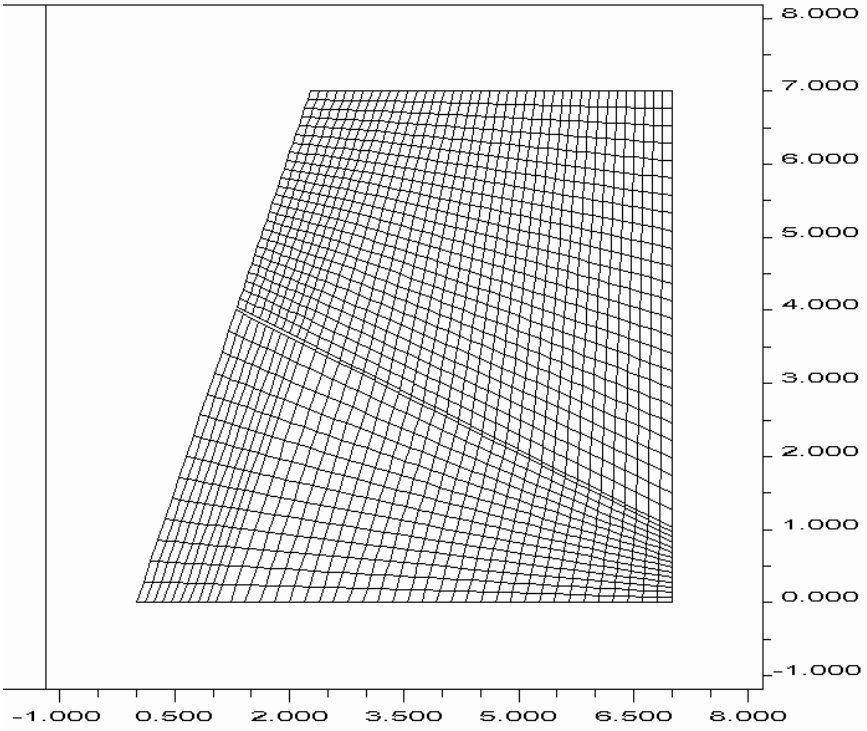


Figure 3.1: Initial Grid

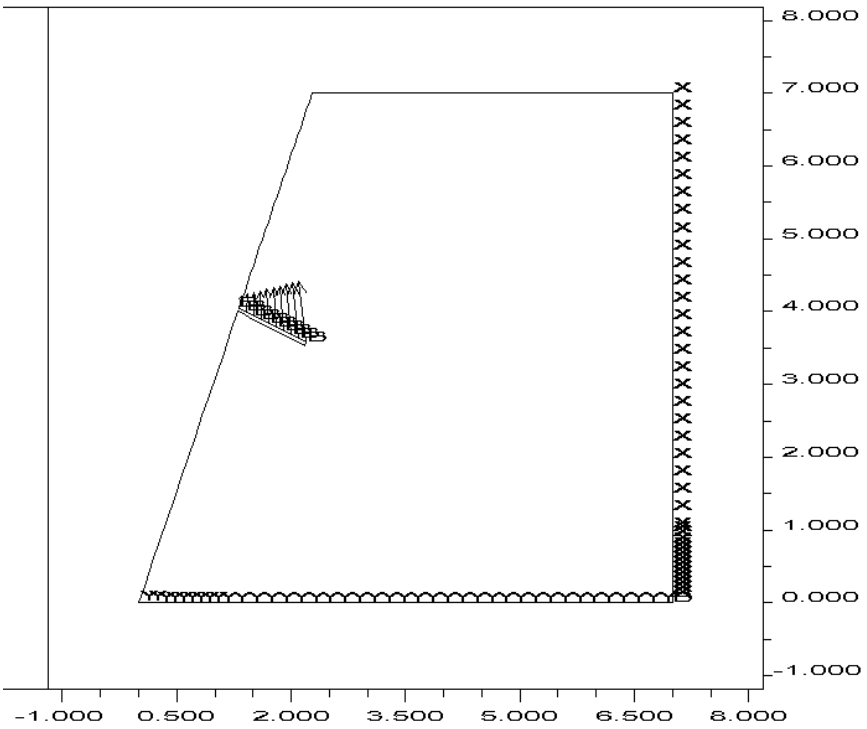


Figure 3.2: Initial Grid with applied forces

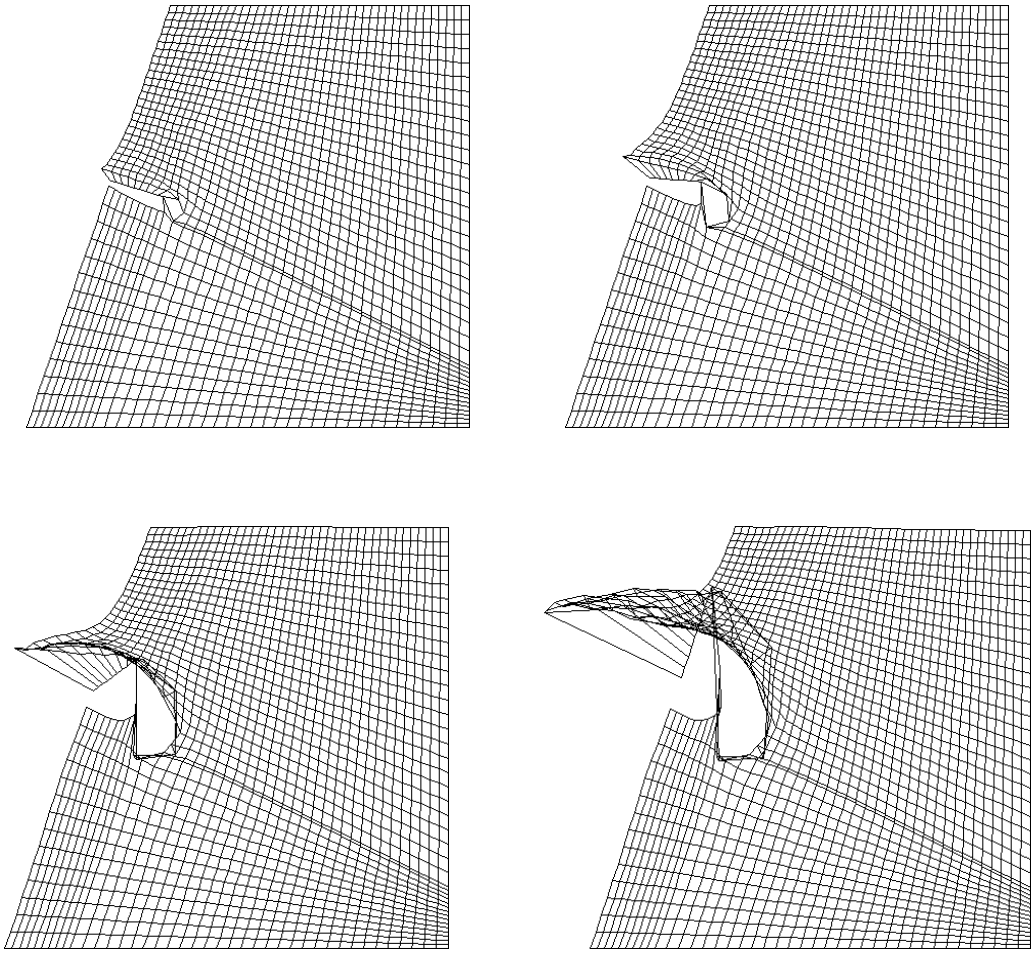


Figure 3.3: Forces after 100, 200, 500 and 1000 steps

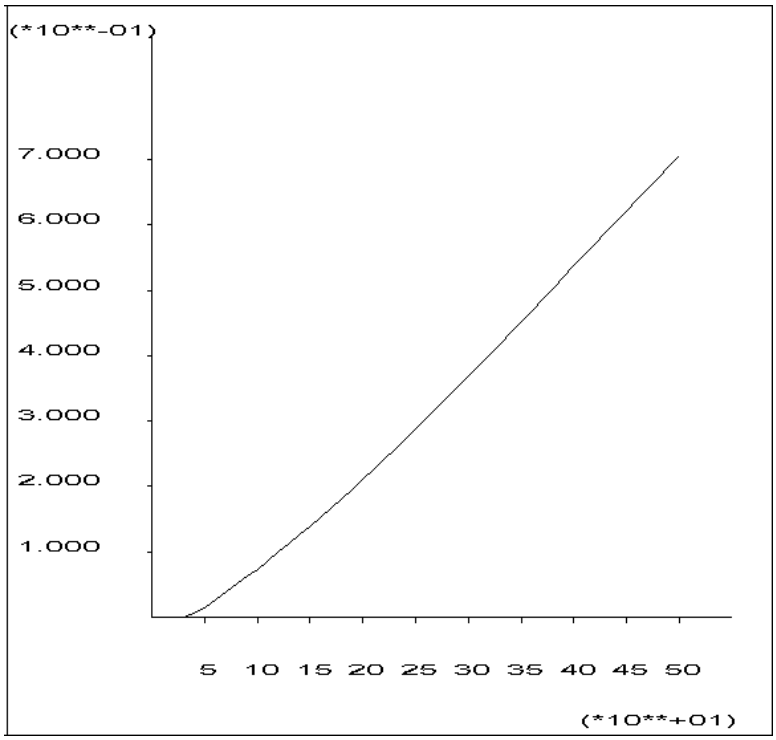


Figure 3.4: Forces after 1000 steps

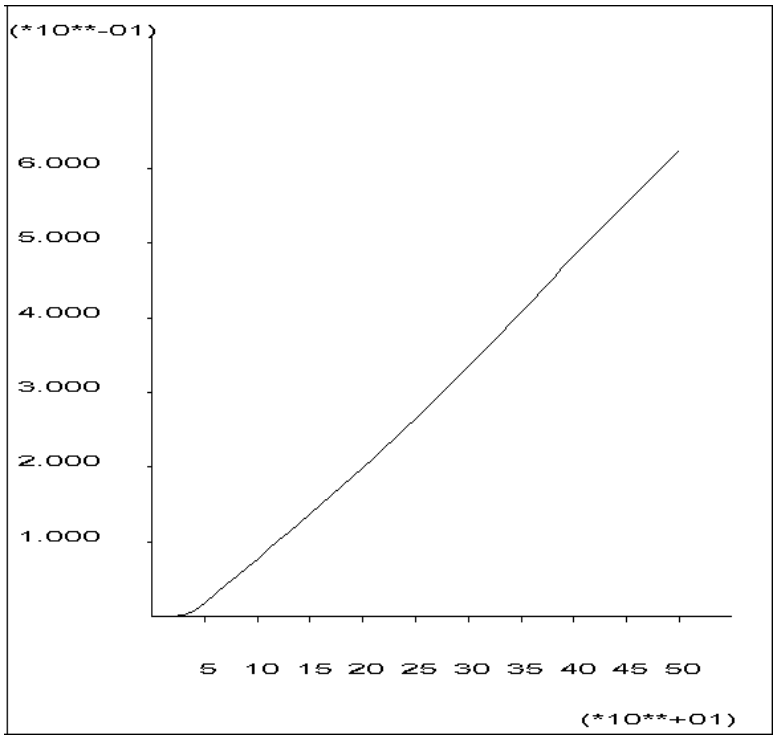


Figure 3.5: Forces after 1000 steps

3.4 Summary

A basic structure for modeling oilsand excavation was implemented using a specialized engineering mechanics software package. We were able to perform some simple investigations on parameter changes to our model. Some methods were discussed in our group which would allow us to determine the volume of oilsands recovered in a single scoop.

In the future, there are commands available in the FLAC program to determine the individual element displacements and area. A count could be performed over the whole grid to determine the total number of elements which were fractured during the scooping process. There are several metrics to determine whether an individual element is fractured. One way is to measure the change in area of the element, and if this change is greater than some fixed value, it will be assumed that there was a fracture within that element.

Many interesting parameters can now be investigated using our model together with some small variations. Various soil properties can be changed to model changes in temperature.

3.5 Sample File Used to Run Simulations

```
;=========================================================================
;--- Oilsands: May 29/99                                               ---
;--- program to calculate shift in sand due to inserted model shovel ---
;=========================================================================

;--- set domain parameters ---
new
def length height theta ybucket phi thickness lengthb
length=7
height=7
theta=18
ybucket=4
phi=10
thickness=0.05
lengthb=1
end

def dev
dev=height*tan(degrad*theta)
end

def xbucket
xbucket=ybucket*tan(degrad*theta)
end

def alpha
alpha=theta+phi
end

def ybuckett
ybuckett=ybucket-(length-xbucket)*tan(degrad*alpha)
end

def gamma
gamma=90-theta-phi
end

def xbucketthick
xbucketthick=xbucket+(thickness*sin(degrad*theta))/cos(degrad*phi)
end

def ybucketthick
ybucketthick=ybucket+(thickness*cos(degrad*theta))/cos(degrad*phi)
end
```

```
def ybuckettt
ybuckettt=ybuckett+thickness/sin(degrad*gamma)
end

def xl
xl=xbucket+lengthb*sin(degrad*gamma)
end

def yl
yl=ybucket-lengthb*cos(degrad*gamma)
end

def xll
xll=xl+(thickness*sin(degrad*theta))/cos(degrad*phi)
end

def yll
yll=yl+(thickness*cos(degrad*theta))/cos(degrad*phi)
end

def xbottom
xbottom=xl-yl/tan(degrad*(90-theta))
end

def xtop
xtop=xbottom+height/tan(degrad*(90-theta))
end

;--- look at calculated parameters ---
print length height theta ybucket phi thickness lengthb
print dev xbucket alpha ybuckett gamma xbucketthick ybucketthick ybuckettt
print xl yl xll yll xbottom xtop

grid 40 40

;--- set up Mohr-Coulomb model and its parameters ---
model mohr
prop b=2e8 c=1e5 d=2e3 di=10 f=10 s=2e8 t=5.67e6
;prop bulk=2e8 shear=2e8 fric=10
;prop dens=2000 coh=0 ten=5.67e6 di=10
m n i=1,9 j=15
;prop bulk=4e10 shear=2e10 dens=2000

;--- ensure bottom and right side of domain does not move ---
;struc beam begin xbucket,ybucket end xl,yl
;struc prop 1 den=2000 he=3
fix y j=1
fix x i=41

;--- set large shows motion of grid element ---
set large
set grav=9.81

gen 0,0 xbucket,ybucket xl,yl xbottom,0 i=1,10 j=1,15
gen xbucketthick,ybucketthick dev,height xtop,height xll,yll i=1,10 j=16,41
gen xll,yll xtop,height length,height length,ybuckettt i=10,41 j=16,41
gen xbottom,0 xl,yl length,ybuckett length,0 i=10,41 j=1,15

;--- various ways to apply forcing at boundary ---
;apply pres=1e6 var 1e7 0 from 1,16 to 10,16
;apply yforce=0 var 1e6 0 i=1,10 j=16
apply yvel=0.0005 var 0.002 0 from 1,16 to 10,16
apply xvel=0.0001 var -0.0004 0 from 1,16 to 10,16
;apply xvel=0.001 from 1,16 to 10,16

plot hold model bou fix apply beam
```

Chapter 4

Print Quality on Paper: Modelling Missing and Perturbed Half-Tone Dot Images

Participants: John Oliver (Mentor), Seema Ali, Ella Huszti, Nathan Krislock, Marni Mishna.

PROBLEM STATEMENT: During printing, statistically, it is possible for ink to transfer to either the paper's pores, the fibrous land areas, or to combinations of these two physical extremes. Consequently, depending on the paper quality, serious limitations may arise in print quality. It would be invaluable to papermakers if, for a pulp of a given average fibre size distribution, they could predict the optimum surface pore size distribution which would have the least effect on the shape and registration of ink dots and hence print quality. Specifically, this project set out to:

- Estimate the statistical probability for microscopic ink drops to miss microscopic land areas, transfer through pores to a sub-layer resulting in mis-register on paper
- Estimate the roundness of microscopic drop images on a porous paper surface with different fibre formation (land/pore distributions).

Chapter 5

Modelling Batch Interfaces

Participants: Don Scott (Mentor), Jesse Bingham, Andriana Davidova, Margaret Liang, Gordon O'Connell, Hope Serate.

PROBLEM STATEMENT: Petroleum products are most economically shipped in long distance pipelines. There can be significant differences between the characteristics of the various petroleum products that are shipped, so they are shipped in batches of fluids with similar characteristics. When different products with different properties (such as density and molecular viscosity) are transported, it is important to be able to distinguish the boundary between them so that upon arrival at the terminal consecutive products can be separated into their respective tanks with a minimum of contamination.

The process of separation is made difficult because during shipping an interface develops between adjacent batches as one component diffuses into the other. This interface degrades the quality of the petroleum product. If it were possible to understand which parameters determine the rate of growth of this interface it may be possible to optimize batch shipments to reduce product degradation.

Here we explore a model of batch interface growth when two fluids flow through a pipe of cylindrical geometry under the driving force of a constant pressure gradient. These fluids may have different viscosity or different densities or both.

5.1 Introduction

Enbridge Pipelines, Inc. presented a problem regarding interface mixing when two different petroleum products are shipped together in a single pipeline. The company maintains a 20 inch diameter pipeline and has observed that the volume of the interface stabilizes at some point during shipping.

Our approach to modeling the interface between two fluids is to present it as a time-dependent diffusion process. Turbulent flow throughout the pipeline ensures homogeneous mixing except at batch interfaces. Once an appropriate solution to the diffusion equation is presented, the remainder of the report describes one way to estimate the volume of the batch interface.

5.2 Model assumptions and fluid properties

Model Assumptions The model described in Section 5.3 makes the following assumptions:

- The pipeline is straight and smooth.
- Pipeline flows are fully turbulent.
- The viscosity and density of adjacent batches in the pipeline are similar. Assume homogeneous mixing.
- The petroleum products in adjacent batches are miscible.

In addition to the model assumptions, the following model parameters are constant:

- $U = 2.235\ m/s$ (flow speed).
- $d = 0.508\ m$ or $0.610\ m$ (pipe diameter).

Fluid Properties Table 5.1 summarizes the fluid properties for a number of petroleum products being shipped in Enbridge's pipeline. The viscosities and densities were provided by Enbridge. This table also contains the Reynolds numbers and coefficients of diffusion for each petroleum product, calculated for 20 inch and 24 inch pipes. The Reynolds number is defined by

$$R = \frac{Ud}{\nu} \tag{5.1}$$

where U is the horizontal velocity of flow in m/s, d is the diameter of the pipe in meters and ν is the viscosity of the fluid in centistokes ($10^{-6}\ m^2/s$).

The coefficient of diffusion was calculated for each petroleum product using the following relation (after Equation 5.1 in [9]):

$$K = 10.1\ a\ U\ \left(\frac{v*}{U}\right) \tag{5.2}$$

where a is the radius of the pipe in meters, U is the velocity of flow in m/s and the quotient $\left(\frac{v*}{U}\right)$ was estimated using the following regression equation

$$\left(\frac{U}{v*}\right) = 5.033602\,(\log_{10}(R)) - 4.013663 \tag{5.3}$$

where R is the Reynolds number estimated using Equation 5.1. This regression was fitted to the data provided in Table 2 in [9]. Notice that the quotient $\left(\frac{v*}{U}\right)$ is completely determined by the Reynolds number.

[1] R_{20} is the Reynolds number for a 20 inch pipe, R_{24} is the Reynolds number for a 24 inch pipe.

[2] K_{20} is the coefficient of diffusion (m^2/s) for a 20 inch pipe, K_{24} is the coefficient diffusion (m^2/s) for a 24 inch pipe. K was estimated using Equation 5.2.

Fluid	**Density** kg/m^3	**Viscosity** centistokes	$R_{20}{}^1$	$K_{20}{}^2$	R_{24}	K_{24}
NGL	560	0.1	11,353,800	0.1820	13,624,560	0.2157
Gasoline	700	0.4	2,838,450	0.2014	3,406,140	0.2383
Diesel Fuel	800	0.9	1,261,533	0.2148	1,513,840	0.2539
Synthetic Crude	585	5.5	206,433	0.2522	247,719	0.2974
Condensate	672	0.5	2,270,760	0.2049	2,724,912	0.2424
Sweet Lights	830	6.0	189,230	0.2543	227,076	0.2998
Sour Lights	852	15.0	75,692	0.2791	90,830	0.3285
Medium Crude	886	55.0	20,643	0.3238	24,772	0.3801
Heavy Crude	920	170.0	6,679	0.3763	8,014	0.4400

Table 5.1: Fluid Properties

5.3 Diffusion model

Suppose that a pipe initially contains fluid A moving at some flow velocity U. At time $t = 0$ fluid B is injected into the pipe at position $X = 0$ immediately adjacent to fluid A. If the flow is turbulent, a region develops along the pipe where fluids A and B mix. This region is called the *interface*.

Assuming a diffusion model, the above situation can be modeled using Fick's second law which is defined by

$$C_t = K\, C_{xx} \tag{5.4}$$

where $C(x,t)$ is the concentration of fluid B in fluid A at time t, x is some axial displacement from the origin $X = 0$ moving in the direction of flow, and, K is the coefficient of diffusion due to the longitudinal components of turbulent velocity [9]. For our purposes the flow velocity, U, is constant. Assume that the concentration of fluid B at the center of the interface is $C = \frac{1}{2}$.

An appropriate solution to this diffusion equation is

$$C(x,t) = \frac{1}{2} - \frac{1}{2} erf\left(\frac{1}{2}\, x\, K^{-\frac{1}{2}} t^{-\frac{1}{2}}\right) \tag{5.5}$$

$$erf(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-y^2}\, dy \tag{5.6}$$

We now show that Equation 5.5 is a solution to Equation 5.4. Let

$$C_t = \frac{\partial}{\partial t}\left[-\frac{1}{\sqrt{\pi}} \int_0^{\frac{1}{2} x K^{-\frac{1}{2}} t^{-\frac{1}{2}}} e^{-y^2}\, dy\right] \tag{5.7}$$

Then by the fundamental theorem of calculus

$$C_t = -\frac{1}{\sqrt{\pi}} \exp\left[-\left(\frac{1}{2}\, x\, K^{-\frac{1}{2}}\, t^{-\frac{1}{2}}\right)^2\right] \frac{1}{2}\, x\, K^{-\frac{1}{2}} \left(-\frac{1}{2}\right)\, t^{-\frac{3}{2}} \tag{5.8}$$

$$= \frac{x}{4K^{\frac{1}{2}}\sqrt{\pi} t^{\frac{3}{2}}} \exp\left(-\frac{x^2}{4Kt}\right), \tag{5.9}$$

$$C_x = -\frac{1}{2}\, \frac{\partial}{\partial x}\left[\frac{2}{\sqrt{\pi}} \int_0^{\frac{1}{2} x K^{-\frac{1}{2}} t^{-\frac{1}{2}}} e^{-y^2}\, dy\right] \tag{5.10}$$

$$= -\frac{1}{\sqrt{\pi}} \exp\left(-\frac{x^2}{4Kt}\right)\, \frac{1}{2}\, K^{-\frac{1}{2}} t^{-\frac{1}{2}} \tag{5.11}$$

$$= -\frac{1}{2K^{\frac{1}{2}} t^{\frac{1}{2}} \sqrt{\pi}} \exp\left(-\frac{x^2}{4Kt}\right), \tag{5.12}$$

and

$$C_{xx} = \frac{x}{4K^{\frac{3}{2}}t^{\frac{3}{2}}\sqrt{\pi}} \exp\left(-\frac{x^2}{4Kt}\right). \tag{5.13}$$

So

$$C_t = K\, C_{xx} \tag{5.14}$$

5.4 Results

5.4.1 Determining the length of the interface

For the model described in Section 5.3 the relation between the length of the interface, S, and the effective diffusivity, K, can be defined as [9, 8]:

$$S = 4\sqrt{Kt}\; erf^{-1}(2y-1) \tag{5.15}$$

where t is the elapsed time of flow from the point of injection to the point of separation (shipping time), K is the coefficient of diffusion and y is the allowable fractional purity. Equation 5.15 suggests that $S \propto t^{\frac{1}{2}}$ and $S \propto K^{\frac{1}{2}}$.

The estimated values of K (Equation 5.2) for several petroleum products is summarized in Table 5.1. Some values for the inverse error function are tabulated in Table 5.2[8].

y	$erf^{-1}(2y-1)$
0.50	0.000
0.80	0.594
0.90	0.901
0.95	1.165
0.97	1.328
0.99	1.640
1.00	—

Table 5.2: Inverse error function at selected values of fractional purity.

Figures 5.1 and 5.2 contain plots of the shipping time versus batch interface volume for NGL, synthetic crude and heavy crude in a 20 inch and 24 inch pipeline respectively.

As an example of how to interpret these plots, consider a scenario where fluid A is synthetic crude and fluid B is NGL. Then by Equation 5.15, if the fractional purity is set to $y = 0.99$, we would expect a 504 m^3 interface to develop between these batches along Enbridge's pipeline between Edmonton and Superior WI assuming a 20 inch pipe and a shipping distance of 1100 miles. For a 24 inch pipe the interface volume would be about 790 m^3. If the relative positions of fluid A and fluid B are reversed (synthetic crude diffusing into NGL), we would expect an interface volume of 593 m^3 for the 20 inch pipe and 928 m^3 for the 24 inch pipe, over the same shipping distance.

5.5 Conclusion

In this report we have studied how the diffusion model described by Taylor [9] might be used to predict batch interface volumes in a bifluid petroleum pipeline. How the predicted interface volumes compare to the observed volumes of contamination is a good subject for future work.

During the analysis of this diffusion model two major concerns arose:

1. The most critical model parameter appears to be K, the coefficient of diffusion. The estimators for K given by Taylor [9] and Sjeitzen [8] are very similar, both authors assuming straight pipelines with smooth walls. Aunicky [1] on the other hand suggests that Taylor's estimator for K is

more appropriate for short haul pipelines. The estimator cited in [1] for long haul pipelines gives estimates two orders of magnitude lower. If this class of diffusion coefficients were used with Equation 5.15 the estimated interface volumes would be significantly lower.

2. The equation used to estimate the volume of batch interfaces (Equation 5.15), contains the parameter y, the fractional purity at the tails of the interface. We found no adequate explanation of how one might use this parameter. Should this value be set high when shipping fluids which are sensitive to contamination? Can this parameter be lowered if a higher degree of contamination is acceptable?

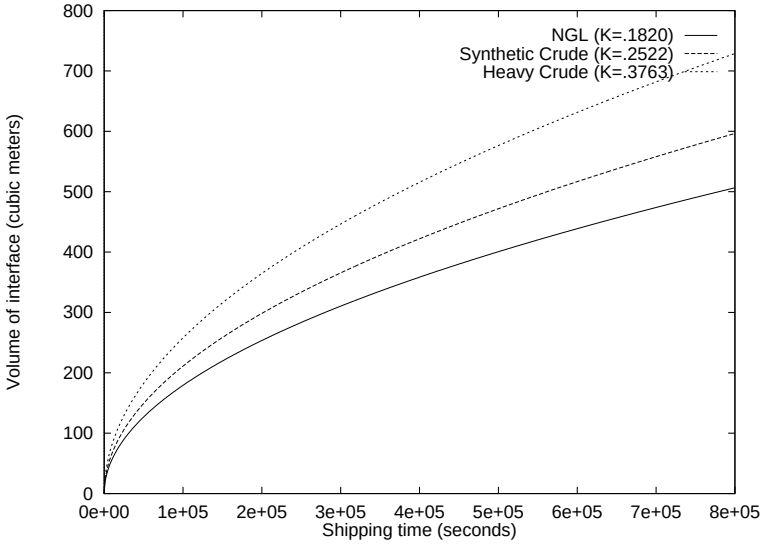


Figure 5.1: Estimated interface volume for a 20 inch pipe.

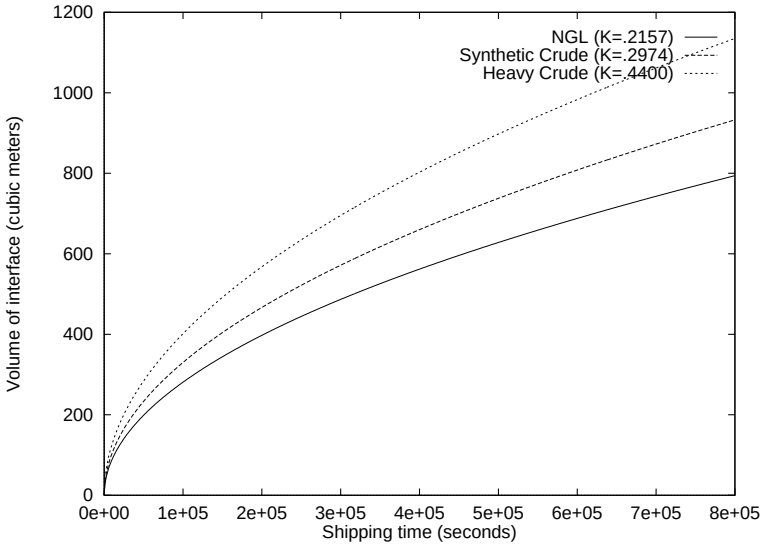


Figure 5.2: Estimated Interface volume for a 24 inch pipe.

Chapter 6

Initiating 3-D Air Target Tracks from 2-D and 1-D Sensor Reports for the Canadian Patrol Frigate (CPF)

Participants: Pierre Valin (Mentor), Hassan Aurag, Vesselin Jungic, Mounia Kijri, Tamara Koziak, Joe Sawada.

PROBLEM STATEMENT: The 2 topics studied here involve target tracking from the Canadian Patrol Frigate (CPF) and can be also used in civilian applications such as airport traffic management. The single target problems presented here are relatively straightforward and will be later used by industry for multiple target problems where association becomes the key issue. The added sensory components furnished by these problems will help resolve the association problem by providing information hidden from the sensors and revealed only through mathematical modeling and related algorithmic processing. In particular, one wishes to initiate a 3-D track from lower-dimensional (2-D for radar and 1-D for Electronic Support Measures) contact data.

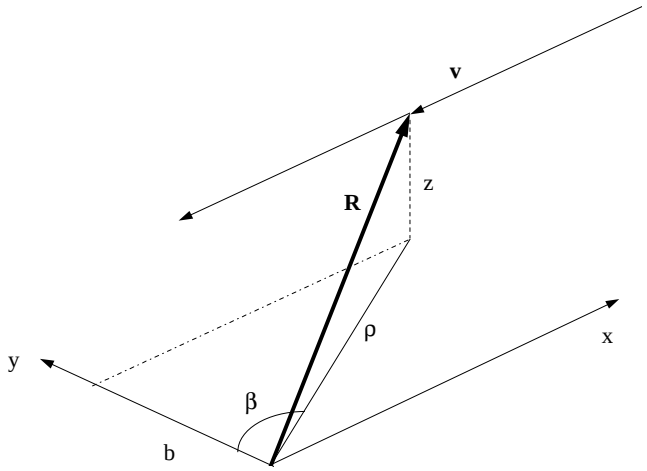


Figure 6.1: The geometry for an aircraft at constant speed and altitude seen by the CPF.

6.1 Introduction

The situation studied is that of an air target being tracked by a 2-D naval radar or a 1-D Electronic Support Measures (ESM) sensor and consists in initiating at long range (or promoting it at a later time) a 3-D track from the 2-D and 1-D sensor reports. The radar is an active sensor that can report slant range and bearing to target, while the ESM is a passive sensor that only detects changes in bearing.

The problem is to find an aircraft's altitude, speed and trajectory (say for a sub-Mach 1 plane) from the CPF's slant range and bearing (w.r.t. true North) report sequence which can be assumed to be at regular intervals (say 5 seconds equivalent to 12 RPM), hereafter referred as contacts (R_i, β_i), $i \in [1, N]$. At large distances the target aircraft will not have moved significantly enough to affect the assumption of nearly regular intervals. Furthermore, we assume that the Earth can be considered flat when the radar range is much smaller than the Earth's radius. The extension to a curved earth involves standard curvilinear coordinates and geodesics. Figure 6.1 shows a simple case of a plane travelling in a straight line at an unknown altitude, an expected behaviour for a plane just entering the radar's outer range extent (and therefore cruising nonchalantly, unaware of the CPF's presence through its own ESM).

Having an answer for the problem above is important when the target has a malfunctioning Interrogative Friend and Foe (IFF) sensor (which can report altitude of a friendly aircraft) or decides to remain circumspect (as would be the case for an enemy aircraft).

Throughout this report, we shall work in the cylindrical coordinates shown in Figure 6.1 and assume that the earth can be considered flat, which is reasonable for a ship with a limited sensor range and thus a limited horizon. This assumption is however not reasonable for an aircraft flying at high altitudes, where the curved earth becomes relevant due to the extended horizon. The cylindrical system centered on the CPF consists therefore of plane polar coordinates in the flat Earth's xy-plane together with altitude forming a righthanded coordinate system. It is related to the traditional Cartesian system (x, y, z) by the convention that the y-axis is in the direction of the true north and that increasing bearing is measured from true north (in a CW direction) as opposed to the traditional method of increasing polar angle from the x-axis in CCW fashion. It is not difficult to see that

$$x = \rho \sin\beta \qquad y = \rho \cos\beta \qquad \vec{r}^{\,2} = R^2 \qquad R^2 = \rho^2 + z^2.$$

We will first describe the simulator from ground truth information from a 2-D radar. This is followed by a description of how this simulated data is used by several cascaded algorithms to correctly extract all the

desired quantities from the target's sensor reports (speed, altitude and trajectory) in the noiseless case. The algorithms are then submitted to noisy data (unclassified sensor accuracies are used). Algorithm design development for 1-D ESM reports then follows. Finally conclusions explain the use of such extracted data for Multi-Sensor Data Fusion (MSDF) and follow-on work is outlined, in particular the need to use classified data for the simulations (especially for the ESM), which is out-of-scope for this unclassified report.

6.2 Radar Data Simulation

When a target is tracked by a naval radar, the sensor reports information regarding the slant range and the bearing (w.r.t. true north) of the target. To simulate this data we create discrete 3-D co-ordinates (x_i, y_i, z_i) to indicate the target's position at time i with respect to the radar sensor. We then use this data to determine the slant range and the bearing. The slant range $R = \sqrt{x^2 + y^2 + z^2}$ and the bearing is computed by $\beta = \arctan(x_i/y_i)$. In the simulated data it is assumed that the radar reports at constant time intervals of 5 seconds (12 RPM). The simulation program was written in C.

In the first simulation we assumed that the radar gave exact information regarding range and bearing, i.e. noiseless data. In practice, however, the data will be perturbed by white noise. We assume that the noise follows a Gaussian distribution. Thus to generate reports with white noise we need to randomly generate Gaussian distributed values to perturb the exact range and bearing. Algorithms to generate normally distributed random numbers with 0 mean and unit variance are given in "Numerical Recipes" by Press et al. Since these numbers are generated with unit variance, we multiply these values by the standard deviation for both the range (100 m) and the bearing (1 degree or 1/57 radians) respectively. The resulting values are then added to the exact range and bearing to simulate reports with white noise. The above numbers are meant only as a guide because in reality radar parameters values tend to be smaller than these values: indeed, for the long range SPS-49, the standard deviations are about half the above quoted values and for the short range SG-150 about 5 times smaller (or better). Throughout this report, representative values are used and the industrial partner will put in the true (often classified) values later.

In one simulation shown in Figure 6.1 we have a target starting at location (in km) (100,100,6) travelling at 0.2 km/s in the negative x direction. If we assume that y is fixed, then we can use the perturbed data to get perturbed values for the positions x and z. We compute x by multiplying the fixed value for y by the tangent of the perturbed angle β_N. For z we use use fixed y, and perturbed x_N and R_N. Thus: $x_N = y \tan \beta_N$ and $z_N = \sqrt{(R_N)^2 - (x_N)^2 - y^2}$.

Note that in this case if the true z is small relative to the actual range R, then by perturbing R even by a small amount we will see a large difference between actual z and z_N. As the target gets closer however (z larger w.r.t. R), this difference decreases as can be seen in Figure 6.2.

6.3 Construction and solution of sets of coupled nonlinear equations for radar reports

Two separate approaches can be investigated, the first can obtain speed and altitude but not the trajectory, while the second can obtain altitude and the trajectory but not the speed. Since all of these quantities are useful, the final solution will consist of a mixture of the two methods described below.

Both the speed and altitude are attributes needed by a Multi-Sensor Data Fusion (MSDF) module for identification. Indeed, one can exclude from all possible air platforms, all those whose documented maximum speed is less than the measured speed and whose documented maximum altitude is less than the measured altitude. The trajectory itself is needed to initialize a Kalman filter with the proper state vector.

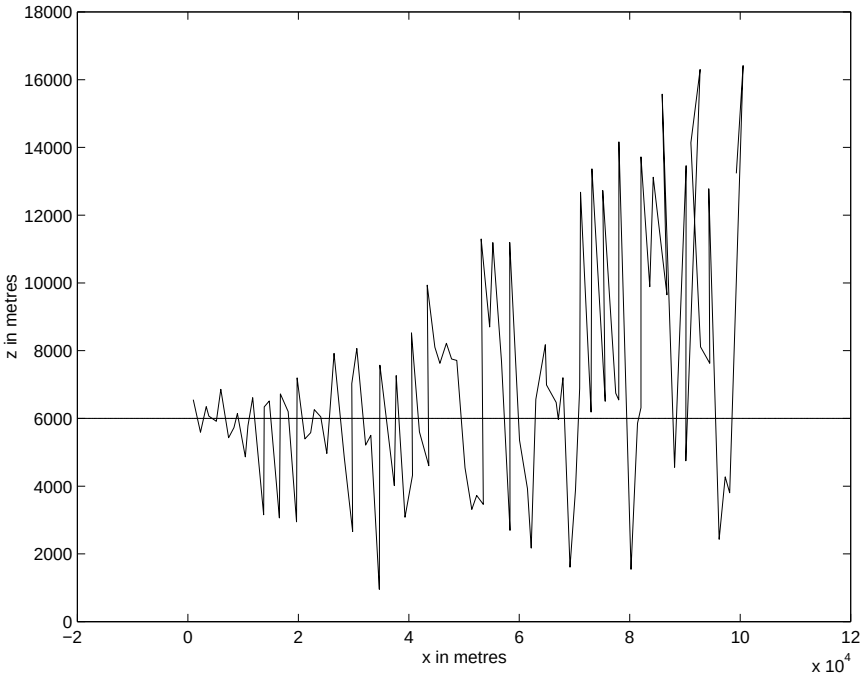


Figure 6.2: Noisy range and bearing data affecting the perceived altitude.

6.3.1 Obtaining speed and altitude

From Figure 6.1, it is easy to see that we have the following relation in the xy plane, since the speed V is constant:

$$|\vec{\rho}_{k+1} - \vec{\rho}_k|^2 = v^2 \Delta t^2$$

Here, $\vec{\rho}_k$ is the projection of the position vector of the plane at contact k onto the xy plane. Explicitly taking the modulus and using $\rho^2 = R^2 - z^2$ we have

$$R_{k+1}^2 + R_k^2 - 2z^2 - 2\sqrt{R_{k+1}^2 - z^2}\sqrt{R_k^2 - z^2}\cos(\beta_{k+1} - \beta_k) = v^2 \Delta t^2$$

For convenience, one can rescale by a fixed value of R (say the first value R_1) and obtain in terms of $\overline{R} = \frac{R}{R_1}$ and $\overline{z} = \frac{z}{R_1}$. This allows dimensionless variables to appear in computer programs, obviating the decision for a system of units.

$$\overline{R}_{k+1}^2 + \overline{R}_k^2 - 2\overline{z}^2 - 2\sqrt{\overline{R}_{k+1}^2 - \overline{z}^2}\sqrt{\overline{R}_k^2 - \overline{z}^2}\cos(\beta_{k+1} - \beta_k) = \left(\frac{v\Delta t}{R_1}\right)^2$$

This equation involves only two unknowns, namely the speed v and altitude z, but does not allow determination of the actual trajectory in the projected $x - y$ plane.

6.3.2 Obtaining the actual trajectory

We represent each point $M'(x, y, 0)$ belonging to the xy-plane in the terms of ρ and β (see Figure 6.3). Here p' is the orthogonal projection of p into the xy-plane.

Note that we assume that the line p is parallel to the xy-plane (at constant z) and that we can measure the values of the range r and the bearing β for a number of points M from the line p. Clearly, the problem is to find the distance between the line p and the xy-plane. Indeed, let the set $\{M_1, M_2, \ldots, M_N\}$

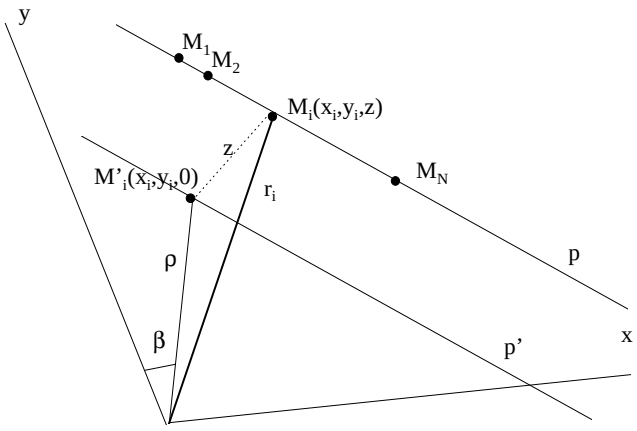


Figure 6.3: Notation for deriving the actual aircraft trajectory.

represents N points of the line p and let the pair (r_i, β_i) represents the range and bearing w.r.t. M_i, for $i \in [1, N]$ (see Figure 6.3). Let P be the plane that contains p and p'. We assume that the line p' does not pass through the origin, i.e., we consider the case where the airplane does not fly overhead. Let $c \in \mathbf{R}$ and let $a, b \in \mathbf{R}$ be such that the equation of P is

$$ax + by = c \ .$$

Hence a and b depend of the choice of c. The idea behind this approach is that with the suitable choice of c we can make our future calculations faster.

From the definition of x and y in terms of ρ and β and the fact that $M_i' \in p'$, we have that for all $i \in [1, N]$

$$\rho_i(a \sin \beta_i + b \cos \beta_i) = c \ .$$

From the triangle OM_iM_i' we have that $R_i^2 - z^2 = \rho_i^2$ and, consequently,

$$(a \sin \beta_i + b \cos \beta_i)^2 (R_i^2 - z^2) = c^2 \ .$$

Let $\overline{z} = \frac{z}{c}$ and $\overline{R}_i = \frac{R_i}{c}$ (for example c can be conveniently chosen to be R_1). Then, for all $i \in [1, N]$

$$(a \sin \beta_i + b \cos \beta_i)^2 (\overline{R}_i^2 - \overline{z}^2) = 1.$$

Hence, the values of $\overline{z}, a$, and b are solutions of this system of N nonlinear equations with 3 unknowns. In the noiseless case, i.e., in the case where the exact values of R_i and β_i, $i \in [1, N]$, are obtained, the system has a solution and that solution is unique.

This approach has three unknowns and experience has shown that any routine that tries to find the simultaneous solution to (at least) 3 such equations is likely to experience problems. In the noisy case, finding the "best fit" to a set of such equations often leads to false minima.

6.3.3 Combined solution

The best solution is obtained by finding v and z through method 1, then inserting that value of z into method 2 to find a and b. Note that although both methods involve equalities in the noiseless case, roundoff errors from the simulator and noisy data in the real-life cases warrant the use of finding only approximate solutions to the sum of squares of such equations. We then have a problem of finding the

χ^2	Fitted z	True z	Fitted v	True v	b/a
4.26992 10^{-9}	4.9708	5.	0.360553	0.3605	-0.666669
4.25659 10^{-9}	7.0107	7.	0.360557	0.3605	-0.666676
4.25963 10^{-9}	8.9964	9.	0.360555	0.3605	-0.666673
4.25678 10^{-9}	11.005	11.	0.360556	0.3605	-0.666675

Table 6.1: Table of results for constant speed for varying altitude.

χ^2	Fitted z	True z	Fitted v	True v	-a/b	True -a/b
4.67839 10^{-7}	5.91902	6.	0.399611	0.4	-0.297673	-0.3
2.43387 10^{-7}	6.07964	6.	0.399998	0.4	0.000028634	0.0
2.1768 10^{-7}	5.9749	6.	0.399604	0.4	0.297641	0.3
4.99305 10^{-8}	5.99207	6.	0.398733	0.4	0.599412	0.6
3.05118 10^{-9}	6.00072	6.	0.399372	0.4	0.898990	0.9
4.94943 10^{-17}	6.	6.	0.398808	0.4	-1.11642	1.0

Table 6.2: Table of results for constant speed and altitude but varying attack angles.

least square solution to an arbitrary number (> 2) of such solutions. We therefore consider the set of the functions

$$g(v,\overline{z}) = \overline{R}_{k+1}^2 + \overline{R}_k^2 - 2\overline{z}^2 - 2\sqrt{\overline{R}_{k+1}^2 - \overline{z}^2}\sqrt{\overline{R}_k^2 - \overline{z}^2}\cos(\beta_{k+1} - \beta_k) - \left(\frac{v\Delta t}{R_1}\right)^2$$

$$f_i(a,b,\overline{z}\ fixed) = (a\sin\beta_k + b\cos\beta_k)^2(\overline{R}_k^2 - \overline{z}^2) - 1, \qquad k \in [1,N]$$

The goal is then to minimize the sum of squares of any of the above equations. One thus starts by minimizing the sum of squares of the first one $min\sum_{k=1}^{N} g_k^2(v,\overline{z})$ to obtain speed and altitude, then inserts the results in the second equation $min\sum_{k=1}^{N} f_k^2(a,b,\overline{z})$ to obtain a and b. Note that results are expected to improve as N increases, particularly for noisy data, since the batch fit includes more points that would fit the straight line. The computation of the least squares solution has been implemented using Mathematica's FindMinimum utility. The usage and implementations notes for the FindMinimum utility are the following. With Method $\rightarrow$ Automatic (which is the one we chose), FindMinimum uses various methods due to Brent: the conjugate gradient in one dimension, and a modification of Powell's method in several dimensions; With Method $\rightarrow$ Newton, FindMinimum uses Newton's method; With Method $\rightarrow$ QuasiNewton FindMinimum uses the BFGS version of the quasi-Newton method.

6.4 Noiseless case for radar contacts

In the first simulation, we take a constant speed, given trajectory and vary the altitude by steps of 2 km. The inputs are various slant range and bearing readings. The output is speed, altitude and direction, as measured by the ratio of b and a.

In the second simulation, we take different angles of approach for the target (or trajectories). The input and output types are the same.

An implementation of such a process should be done in C. Such a C-routine should first start by getting a good speed evaluation. With the speed obtained after N-steps at an acceptable precision, the routine should then restart by setting the speed correctly. It then should discard any off-range

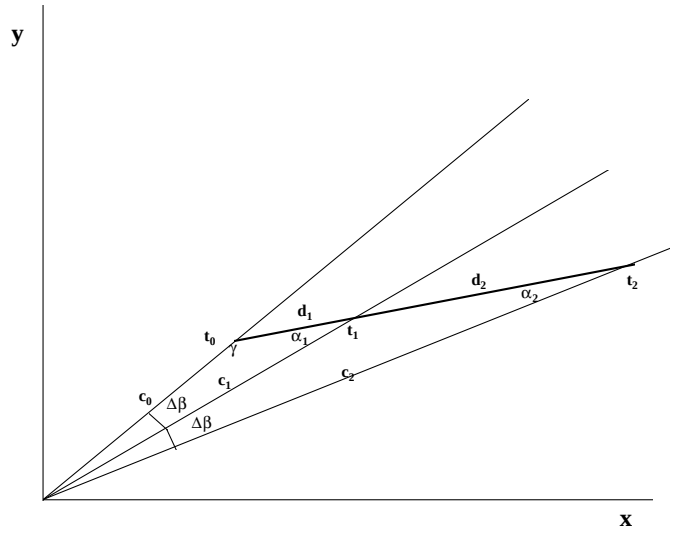


Figure 6.4: ESM geometry construction in the $x - y$ plane.

readings of slant ranges. The previously shown Figure 6.2 shows how noise affects readings of the altitude and explains the failure of the straightforward approach to compute the altitude. One should also recall, that we do obtain the correct speed even in noisy environment. Benchmarking can be done through simulated noiseless data versus noisy data. The bias in noisy data can be tabulated and the outliers removed through these simulations. Real data, being intrinsically noisy, would then use the same outlier truncation mechanism. One also probably will have to use something different from least squares minimization. An example would be using maximum logarithm likelihood.

6.5 ESM Bearing-only Reports - two approaches

The sensor being used to track targets on Electronic Support Measures unit (ESM) only reports a time-sequence of their bearing $\beta(t_i)$ when the bearing change attains a certain value $\Delta\beta$. This naval sensor cannot report the airborne target's slant range nor its projected range on the Earth. Because of the classified nature of the ESM parameters $\Delta\beta$ and its maximum range, no simulations were performed at the Workshop and the methods will be tested for effectiveness by the industrial partner at a later time. The problem that we discuss in this section is how one can obtain a rough estimate of the target's slant range (and ultimately its range and altitude, by using the result s of Section 6.3.1) from the observed bearing rate (i.e., a sequence of different t_i for a constant $\Delta\beta$) through analytical and numerical analysis/simulations. Again there are two competing approaches which have the potential to be combined into a final optimized solution.

6.5.1 One Approach

Figure 6.4 shows the projection on the plane xOy of the trajectory of the aircraft.

The purpose of this part is to be able to determine the direction of the aircraft. We must keep in mind that we suppose that the velocity v of the aircraft is constant. If Δt_1 is the variation of time between t_0 and t_1, and d_1 is the associate displacement, then we have

$$d_1 = v\Delta t_1.$$

Similarly, if Δt_k is the variation of time between t_{k-1} and t_k, and d_k is the associate displacement, then we have

$$d_k = v\Delta t_k.$$

So using those two equations, we get

$$d_k = \frac{\Delta t_k}{\Delta t_1} d_1.$$

Now, using the law of sines for the two triangles obtained in the case where $k = 1$ and $k = 2$, we get two equations:

$$\frac{\sin \Delta\beta}{d_1} = \frac{\sin \alpha_1}{c_0}$$

and

$$\frac{\sin 2\Delta\beta}{d_1 + d_2} = \frac{\sin \alpha_2}{c_0}.$$

Since $\Delta\beta + \alpha_1 + \gamma = \pi$ and $2\Delta\beta + \alpha_2 + \gamma = \pi$, we get $\alpha_2 = \alpha_1 - \Delta\beta$. On the other hand, we also have $d_2 = \frac{\Delta t_2}{\Delta t_1} d_1$. So the two equations become:

$$\frac{\sin \alpha_1}{\sin \Delta\beta} = \frac{c_0}{d_1}$$

and

$$(1 + \frac{\Delta t_2}{\Delta t_1}) \frac{\sin(\alpha_1 - \Delta\beta)}{\sin 2\Delta\beta} = \frac{c_0}{d_1}.$$

We finally have:

$$\sin \alpha_1 = (1 + \frac{\Delta t_2}{\Delta t_1}) \frac{\sin(\alpha_1 - \Delta\beta)}{2 \cos \Delta\beta}.$$

We then obtained α_1, and since $\Delta\beta + \alpha_1 + \gamma = \pi$, we get γ, and so we have the direction of the plane. If we go back to the law of sines, we have

$$\frac{\sin \Delta\beta}{d_1} = \frac{\sin \alpha_1}{c_0}$$

and

$$\frac{\sin k\Delta\beta}{d_1(1 + \sum_2^k \frac{\Delta t_i}{\Delta t_1})} = \frac{\sin \gamma}{c_k}.$$

So if we just have c_0, we then get d_1 (so we know the plane's speed), and so all the c_k. We then have the trajectory of the plane and not only its direction. This could be achieved for example by using only a few radar returns (ensuring near electromagnetic silence of the CPF) or by using Link-11 time-updated tracks from Participating Units for an approximate range.

6.5.2 Another Approach

Suppose that the velocity of the target is given by

$$\vec{v} = v_x \vec{i} + v_y \vec{j}\ .$$

Also suppose that N contacts have been reported. Let $\{M_1, M_2, \ldots, M_n\}$ is the set of points so that the point M_n represents the position of the target at the time t_n of the nth contact. Let $\Delta t_n = t_{n+1} - t_n$ and let $\beta_n = \beta(t_n)$ be the bearing reported at the time t_n. We assume that $\beta_1, \ldots, \beta_N$ form an arithmetic progression with the common difference equal $\Delta\beta$, where $\Delta\beta$ is positive but very small. As before, r_n is the slant range of M_n and ρ_n is its projected range. Figure 6.5 illustrates the situation we shall discuss.

Since the velocity is $\vec{v} = v_x \vec{i} + v_y \vec{j}$, we have that

$$x_{n+1} = x_n + v_x \cdot \Delta t_n$$
$$y_{n+1} = y_n + v_y \cdot \Delta t_n$$

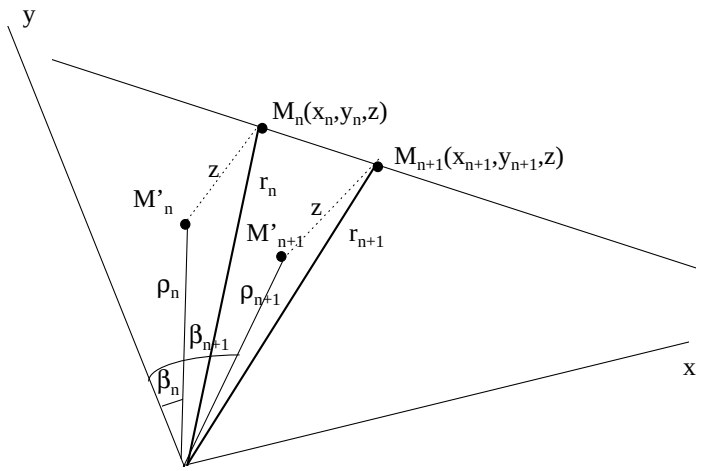


Figure 6.5: ESM geometry construction.

such that, as before,

$$\rho_{n+1}\sin\beta_{n+1} - \rho_n \sin\beta_n = v_x \cdot \Delta t_n$$
$$\rho_{n+1}\cos\beta_{n+1} - \rho_n \cos\beta_n = v_y \cdot \Delta t_n \ .$$

We consider these two equations as a linear system with ρ_{n+1} and ρ_n as unknowns. Note that

$$\begin{vmatrix} \sin\beta_{n+1} & -\sin\beta_n \\ \cos\beta_{n+1} & -\cos\beta_n \end{vmatrix} = -\sin\Delta\beta \neq 0 \ .$$

Thus,

$$\rho_n = \frac{v_x\cos\beta_{n+1} - v_y\sin\beta_{n+1}}{\sin\Delta\beta}\cdot\Delta t_n \ .$$

Consequently, for any $n \in [1, N]$

$$z^2 = R_n^2 - \frac{(v_x\cos\beta_{n+1} - v_y\sin\beta_{n+1})^2}{(\sin\Delta\beta)^2}\cdot(\Delta t_n)^2 \ .$$

Hence, from the given data, we have obtained a nonlinear system with N equations and $N+3$ unknowns $z, v_x, v_y, R_1, \ldots, R_n$ One observes that the projections of $M_1, \cdots, M_N$ into xy plane belong to the line

$$v_y x - v_x y = c$$

where c is such that

$$c = v_y\rho_1\sin\beta_1 - v_x\rho_1\cos\beta_1 = \frac{(v_x\cos\beta_2 - v_y\sin\beta_2)(v_y\sin\beta_1 - v_x\cos\beta_1)}{\sin\Delta\beta}\cdot\Delta t_1 \ .$$

Thus, the line p is given, as before, by $Ax + By = C$ where

$$A = \frac{v_y}{(v_x\cos\beta_2 - v_y\sin\beta_2)(v_y\sin\beta_1 - v_x\cos\beta_1)} \ ,$$

$$B = -\frac{v_x}{(v_x\cos\beta_2 - v_y\sin\beta_2)(v_y\sin\beta_1 - v_x\cos\beta_1)} \ ,$$

and

$$C = \frac{\Delta t_1}{\sin \Delta\beta} \ .$$

To obtain an approximate value of z we follow these steps.

1. Choose $\zeta \in \mathbf{R}^+$.
2. Define, for each $n \in [1, N]$

$$r_{n,\zeta} = r_{n,\zeta}(v_x, v_y) = \sqrt{\zeta^2 + \frac{(v_x \cos\beta_{n+1} - v_y \sin\beta_{n+1})^2}{(\sin\Delta\beta)^2} \cdot (\Delta t_n)^2} \ .$$

3. Solve the system

$$(A \sin\beta_n + B \cos\beta_n)^2 (\overline{R}^2_{n,\zeta} - \overline{z}^2) = 1$$

where

$$\overline{R}_{n,\zeta} = \frac{R_{n,\zeta}}{C} \ .$$

Note that the unknowns are z, v_x, and v_y. Also, this situation is the same as the problem that we discussed in Section 6.3.1.

Alternatively, find $(\overline{z}_\zeta, v_{x,\zeta}, v_{y,\zeta}) \in \mathbf{R}^3$ so that if

$$g_i(z, v_x, v_y) = (A \sin\beta_n + B \cos\beta_n)^2 (\overline{R}^2_{n,\zeta} - \overline{z}^2) - 1$$

then

$$\sum_{i=1}^{N} (g_i(z_\zeta, v_{x,\zeta}, v_{y,\zeta}))^2 = \min\{\sum_{i=1}^{N} (g_i(z, v_x, v_y))^2 : (z, v_x, v_y) \in \mathbf{R}^2\}.$$

4. If $(\overline{z}_\zeta, v_{x,\zeta}, v_{y,\zeta})$ is the solution of the system from 3. above, find

$$z_\zeta = \overline{z}_\zeta \cdot C.$$

5. Compare z_ζ and ζ. If z_ζ and ζ are "close", we are done. If they are not, iterate over ζ, which describes a family of possible solutions, until z_ζ and ζ are "close" enough.

6.6 Conclusions and Outlook

This project studied how to initiate (or promote) higher dimensional air tracks from lower-dimensional contacts. At least two design solutions were constructed for each of the two cases: a 2-D radar and an ESM sensor. The complete solution was implemented for the 2-D radar with perfect results in the noiseless case and promising results for the noisy case. Due to the classified nature of ESM specifications, the ESM designs will be implemented at a later date by the industrial partner.

The techniques used in this report determine first the speed and altitude which are attributes needed by a Multi-Sensor Data Fusion (MSDF) module for identification. Indeed, one can exclude from all possible air platforms all those whose documented maximum speed is less than the measured speed, and whose documented maximum altitude is less than the measured altitude. These values have been tabulated by Lockheed Martin (LM) Canada in a Platform Database which now contains over 140 platforms. For each of these deduced attributes, propositions can be constructed that contain all platforms with consistent attributes and these propositions are then fused in an evidential reasoning scheme such as a truncated Demspter-Shafer algorithm developed at LM Canada. Finally, the trajectory itself is needed to initialize a Kalman filter with the proper state vector. In addition, optimization routines can provide an estimate of parameter errors (here velocity and altitude, as well as the actual trajectory), which can be used to construct an initial covariance matrix for the Kalman filter.

Bibliography

[1] Z. Aunicky. The longitudinal mixing of liquids flowing successively in pipelines. *The Canadian Journal of Chemical Engineering*, 48:12–16, February 1970.

[2] C. M. Bethke and Marshak. Brine migrations across north america—the plate tectonics of groundwater. *Annual Reviews Earth and Planetary Sciences*, 18:287–315, 1990.

[3] G. Dagan. Statistical theory of groundwater flow and transport: Pore to laboratory, laboratory to formation, and formation to regional scale. *Water Resources Research*, 22:120–135, 1986.

[4] G. de Marsily. *Quantitative Hydrogeology*. Academic Press, San Diego, CA, 1986. pp. 13–22, 39–40, 50, 58–64, 80–82, 284–329.

[5] A. Desbarats and S. Bachu. Geostatistical analysis of aquifer heterogeneity from the core scale to the basin scale: A case study. *Water Resources Research*, 30:673–684, 1994.

[6] Donald E. Hull and James W. Kent. Radioactive tracers to mark interfaces and measure intermixing in pipelines. *Industrial and Engineering Chemistry*, 44(11):2745–2750, November 1952.

[7] W. W. Rubey and M. K. Hubbert. Role of fluid pressure in mechanics of overthrust faulting. *Geological Society of America Bulletin*, 70:167–206, 1959.

[8] F. Sjenitzer. How much do products mix in a pipeline? *The Pipeline Engineer*, pages D31–D34, December 1958.

[9] Geoffrey Taylor. The dispersion of matter in turbulent flow through a pipe. *Proceedings of the Royal Society of London*, 223 A:446–468, 1954.

List of Participants

Organising Committee

Abel Cadenillasi	University of Alberta
Douglas Kelker	University of Alberta
Mike Kouritzin	University of Alberta
Henry Leung	University of Calgary
Bryant Moodie	University of Alberta
Bruce Sutherland	University of Alberta
Yanhong Wu	University of Alberta

Mentors

Stefan Bachu	Alberta Energy Utilities Board
Wayne Grover	TRLabs Ltd.
Mike Lipsett	Syncrude Ltd.
John Oliver	Alberta Research Council
Don Scott	Enbridge Pipelines Ltd.
Pierre Valin	Lockheed Martin Canada Ltd.

Students

Tamar Dakic	Simon Fraser University
Marni Mishna	Simon Fraser University
Vesselin Jungic	Simon Fraser University
Andrea Amariel	University of Alberta
Brenda Hawkins	University of Alberta
Todd Oliynyk	University of Alberta
Paul Wiebe	University of Alberta
Ella Huszti	University of Alberta
Tamara Koziak	University of Alberta
Shelly Pinder	University of British Columbia
Scott MacLachlan	University of British Columbia
Margaret Liang	University of British Columbia
Luz Palacios	University of Calgary
Hope Serate	University of Calgary
Devom McCrea	University of Victoria
Peilin Shi	University of Victoria
Seema Ali	University of Victoria
Jesse Bingham	University of Victoria
Andriana Davidova	University of Victoria
Gordon O'Connell	University of Victoria
Joe Sawada	University of Victoria
Hassan Aurag	University of Montreal
Mounia Kijri	University of Montreal
Vincent Lemaire	University of Montreal
Emmanuel Ngimbi Ngembo	University of Montreal
Nathan Krislock	University of Regina
Grace So	University of Toronto

PIMS Contact Information

email: pims@pims.math.ca
http://www.pims.math.ca

- **Director: N. Ghoussoub**
  Admin. Asst: J. Burian
  Phone: 604-822-3922
  Fax: 604-822-0883
  email directorpims.math.ca
- **SFU-Site Director: R. Russell**
  email: sfu@pims.math.ca
- **UAlberta-Site Director: B. Moodie**
  email: ua@pims.math.ca
- **UBC-Site Director: D. Rolfsen**
  email: ubc@pims.math.ca
- **UCalgary-Site Director: M. Lamoureux**
  email: uc@pims.math.ca
- **UVic-Site Director: F. Diacu**
  email: uvic@pims.math.ca
- **UWashington-Site Director: T. Toro**
  email: uw@pims.math.ca

Editor: Heather Jenkins, PIMS