Project Summaries

Projects are the backbone of CQADS’s activities: in the two short years since its inception, the Centre has tackled no less than 15 projects, covering nearly all aspects of quantitative analysis. Prior to joining CQADS, its consultants have also been involved in a variety of academic and industry studies.

Three projects were selected to showcase CQADS’ experience conducting analytical and quantitative research and in succesfully implementing advanced theoretical fields to obtain tangible and useful results:

1. Wait-Time Impact Model at Pre-Board Screening Checkpoints for Canadian Airports

Investigators: **Zhao, Y.**, **Boily, P.**, Ye, W.

Client: Canadian Air Transport Security Authority

Domain: **Airport Security Environment**

Subject Areas: **Stochastic Modeling and Analyses**, **Operations Research** and Optimization

1. Analysis of Fluidity Indicators and Seasonality Adjustments for Containers Transit Times in a Multi-Modal Supply Chain Networks

Investigators: **Boily, P.**, Huang, Y.

Client: Transport Canada

Domain: **Transportation**

Subject Areas: **Statistical Analysis and Estimation**, **Forecasting** and Predictive Analytics

1. Discrimination and Non-Discrimination in Parasites: A Simulation Created for Analysis of Evolutionarily Stable States and Population Dynamics

Investigator: **Schellinck, J.**

Client: Forbes Laboratory

Domain: Population Dynamics

Subject Area: **Simulations**

While additional projects could have been brought forward to guarantee that all project summaries would be related to transportation, aviation, and airport security environment, no combination of three projects would have successfully managed to demonstrate CQADS’ experience in the provision of services in all 5 Subject Areas. In the same vein, this approach allows us to showcase a wider (albeit still incomplete) range of proposed team members.

Other projects could have been selected to showcase our expertise with Predictive Analytics (Classification Trees and Random Forests for Public Health Agency of Canada; Clustering Analysis for United Way Centraide Canada), Optimization (Combinatorial Clock Auction for Nordicity Group Ltd.; Scheduling Algorithm for BluInk) or another transportation-flavored project (the award-winning Canadian Vehicle Use Study for Transport Canada); summaries are available upon request.

**Note:** the first two projects summaries are intended to highlight the relevance and similarity in scope of typical CQADS projects, as well as in their analytical challenges and their complexity, with potential CATSA projects as specified in the RFSO.

The third project summary places greater emphasis on the project processes and difficulties, as well as its results.

## Project Summary – CATSA

**Wait Time Impact Model at Pre-Board Screening Checkpoints for Canadian Airports**

by Yiqiang Zhao, Patrick Boily, Wenzhe Ye

Centre for Quantitative Analysis and Decision Support, Carleton University

**CLIENT ORGANIZATION**

The Canadian Air Transport Security Authority (CATSA) ensures the safety and well-being of passengers as they board their flights at Canadian airports each year. A federal crown corporation founded in the aftermath of the September 11, 2001 terrorist attacks on American soil, CATSA protects the public by efficiently screening air travellers and their baggage at designated airports.

CATSA offers a number of security-related services including pre-board screening, where passengers and their belongings are searched for prohibited and potentially dangerous items such as knives, firearms and explosives. CATSA also conducts hold-baggage screening – where checked baggage is screened using explosive detection equipment – and non-passenger screening where airport workers who have access to restricted areas are searched. Finally, CATSA also administers and maintains the Restricted Area Identity Card (RAIC) program.

**PROJECT INTENT, SCOPE, AND OBJECTIVES**

Numerous factors influence the wait time at pre-board screening checkpoints at Canadian airports: the schedule intensity of departing flights, the volume of passengers on these flights, the number of servers and processing rates at a given checkpoint, etc.

One of CATSA's goals is to ensure that the pre-board screening experience at Canadian airports is made as efficient as possible by minimizing the waiting time at checkpoints. In order to help CATSA gain a better understanding of the waiting process, CQADS developed a Wait Time Impact (WTIM) model using a Queueing Theory approach.

The original scope of the project consisted of:

1. Provide estimates of the passenger arrival rates , the processing rates and the number of servers at each checkpoints, using the available field data
2. Calculate the Quality of Service (QoS) level and determine what service level can be achieved at each checkpoint (i.e. the percentage of passengers which will wait less than minutes, for fixed) for a given arrival rate , processing rate and number of servers .

Provide the average number of servers required to achieve a prescribed QoS level , given an arrival profile .

Implement the WTIM on a SAS platform to allow for the analysis of various scenarios (such as passenger growth, for instance) via the tweaking of a small number of parameters and whenever the available data is updated.

Upon satisfactory completion of these objectives, a second phase was initiated at CATSA’s behest, with the intent of extending their implementation of the WTIM. This second phase’s scope consisted of three objectives:

1. Provide quality of service (QoS) level curves (i.e. cumulative distribution curves) under various arrival rate and number of active servers for each checkpoint (where is allowed to vary).
2. Seamlessly integrate the WTIM with CATSA’s scheduling optimizer, in order to implement a one-click SAS program.
3. Provide validation and modeling analytics support for the integrated WTIM.

**METHODOLOGY**

In order to complete the assignment, CQADS used the following methodological steps:

1. *Exploration of available data*, in order to identify any underlying patterns and essential characteristics.
2. *Understanding the conceptual model*: including document review pertaining to CATSA’s existing framework to gain a full understanding of the structure of its queueing system.
3. *Estimation of model parameters*, which required: making appropriate assumptions to simulate the processes in the queueing system according to the knowledge gained through data exploration; selecting appropriate parameter estimation methods, using the appropriate statistical inference and/or numerical method, based on the completeness and characteristics of the existing data; and conducting parameter estimation accordingly.
4. *Implementation of the conceptual model* on a SAS platform, which allowed for the discovery of the importance of certain notions whose importance only emerged after running some early scenarios through the modification of a small number of parameters (arrival profile, service time distribution, number of servers, service level, etc.), in particular when it came to vacation policy regarding the number of lines, which lead to a switch to a generalized model.
5. *Validation of the generalized model*, by comparing the estimated characteristics of the prototype queueing model (e.g. inter-arrival and service time distributions, average idle time per server, etc.) with their empirical counterparts to determine the validity of the conceptual model. The conceptual model was found to be mostly invalid until a key link between the average arrival rate, the processing rate and the number of lines was established. This combined generalized and Regression model produced good results in most cases, but in certain instances, a departure from the empirical data could still be identified. Further analysis lead to a breakthrough and the introduction of a Departure parameter. The final model, then, combined the , Regression and Departure hypotheses.
6. *Performance evaluation of the final model* was achieved in two ways. A preliminary performance evaluation pitted the model favourably against historical data, but the ultimate test came once predictions were compared to data that were collected after the final model was delivered, again very favourably.
7. *Documentation of the final model*: a technical report providing an overview of the model, as well as describing and justifying the various assumptions, was written and delivered to CATSA stakeholders.
8. *Knowledge transfer* was achieved through meetings (in person or by phone) and email exchanges detailing the progress, increasing in frequency as the deadline approached.
9. *Provision of on-going support* to CATSA’s model users allowed for a number of improvements, both in scope and in implementation.

**PROJECT SUMMARY**

The available data covered 26 checkpoints, at 8 Canadian airports. At each checkpoint, the pre-board screening process is structurally similar: passengers arriving at the beginning of the main queue may have their boarding passes scanned at the position (the start of the waiting queue), but they are always scanned at the position (as they are being processed).

For each checkpoint, 3 datasets were available for each year:

* the *Raw Data* which contains – for each passenger reaching the end of the queue at – the date, scan time at , scan time at and the wait time between and ;
* the *Checkpoint Utilization Report* which records – for each day of the year and each non-overlapping 15-minute block – the maximum number of open processing lines, and
* the *Waiting Time Report* which consists of the subset of the *Raw Data* for which and are both available (and for which observations with anomalous and/or outlying wait time behaviour have been removed by CATSA).

The data was then grouped into meaningful clusters exhibiting properties that can be characterized by the same Poisson process, which allows for proper estimation of queueing model parameters, under the assumption that the queueing model model was valid.

[One difficulty with this approach is that, in practice, the number of servers varies with time, according to a vacation policy which depends on a variety of factors. As such, it is extremely difficult to model. This is problematic since the sought QoS level depends not only on the arrival rates, but also on the processing rates, which themselves depend, among other things, on the number of open servers. Switching to a generalized server (behind which the actual servers are hidden) circumvents this issue, but at the cost of not immediately being able to retrieve the number of servers from the generalized model.]

The average arrival rates for each cluster were computed from the *Raw Data* using Burke’s Theorem and were shown to indeed follow a Poisson process as the inter-arrival times between consecutive events were i.i.d. exponential random variables with parameters lending support to the generalized hypothesis. The average wait times were then estimated using the *Wait Time Report*.

[An analysis of the reasons for the omission of those observations without an scan from the *Wait Time Report* suggests that using the latter to estimate the cluster average wait times is likely to affect the predicted QoS levels, especially in the small wait time regime. However, since short wait times are not likely to cause consternation among the general public, this issue may not arise in practice and can be side-stepped in the estimation phase.]

The estimated processing rates and QoS levels were easily recovered from the relationships

where represents the estimated traffic intensity.

Since these relations do not hold if the generalized hypothesis fails, the need to validate it became more pressing. The simplest way to do so was to compare the wait times generated by the model to those of the empirical data: were the estimated QoS curves “close to” the empirical QoS curves Using two different metrics (largest relative difference ratio, largest area ratio), we showed that the generalized assumption, while not exact, is a reasonable one to make at the checkpoint level.

[This result is achieved without explicitly invoking the number of open servers . Granted, that number is implicitly involved in the determination of the average wait times , but it does not change the fact that it cannot be recovered using solely the tools provided by queueing theory. The Regression assumption asserts that, on a quarterly level, the cluster processing rates is a function of the number of active servers (hidden behind the generalized server) and the arrival rates and that this functional relationship is the same for all regression clusters making-up a given quarter.]

Using the *Checkpoint Utilization Report*, the average service rates per line and average arrival rates per line were estimated for each checkpoint, quarter, and cluster, and then regressed against one another to determine the optimal regression parameters , yielding new estimates for the cluster processing rates. Thus, estimates for the QoS level were easily computed, without explicitly referring to processing rates, using

which held as a direct consequence of the combined and Regression assumptions.

Using the two validation metrics introduced above, it was shown that the combined assumptions, while proving slightly less valid than the hypothesis on its own, still provided reasonably close QoS estimates at the quarter and checkpoint levels.

[This lessened validity should come as no surprise, as there is no way to extract the number of clusters without postulating an external relationship of the form , and that the simple linear regression form used necessarily introduces some uncertainty. Some of that uncertainty might decrease using a more complex regression function.]

In order to predict the number of servers required to meet a given QoS level at a given checkpoint during a given quarter (i.e. for a given pair of regression parameters ), for a given arrival profile , it then sufficed to solve for , yielding

where is the main branch of the Lambert -function.

[Unfortunately, cannot be evaluated by elementary needs except at special values and so one has to depend on efficient numerical algorithms to recover . As SAS does not lend itself particularly well to repeated algorithmic computations, it becomes imperative to find a quick and relatively accurate alternative approach, such as the following approximation, which can be implemented in SAS:

valid for ]

For any given checkpoint, quarter, and cluster, it was thus possible to compare the actual number of open servers (given by the *Checkpoint Utilization Report*), and the estimated value given the actual arrival rate and the actual QoS level Plotting against for all clusters strongly suggested that the prediction and the actual values were linked at the checkpoint level according to for some checkpoint departure parameter . Computed values of near 1 for nearly all checkpoints further validated the combined model. The final prediction for the number of servers was further refined by setting

In theory, it is thus possible to forecast the number of servers for a cluster using only its regression parameters , its departure parameter , an arrival rate , and a QoS level . The validation procedure in this case is slightly different: it makes little sense to compare the predicted value with the actual number of servers found in the historical data as the prediction depends not only on the forecasted arrival rate (which is likely to be different from the historical rate), but also on the attained QoS level (for which an independent forecast is unavailable).

The best validation alternative, then, is to wait for new data to be collected, to determine the actual cluster arrival rate and QoS level to be used in the forecast in order to provide a prediction , and to compare it with the actual recorded over the data collection period.

**DIFFICULTIES AND ADDITIONAL TASKS**

Apart from the usual issues surrounding the transfer of large datasets, the specific issue of the lack of algorithmic computability of the Lambert -function in SAS, and the technically difficult modeling problem, additional tasks were requested by CATSA and the scope of the project was accordingly extended.

These tasks included:

* the modification of the originally requested queueing model to a queueing model;
* the introduction of the combined and regression hypothesis to recover the number of servers ;
* the identification and clean-up of data integrity issues for three airports (YUL, YYC, YVR), and
* the conversion of a MATLAB non-linear solver for the computation of the Lambert -function to an approximation which could be implemented in SAS.

**RESULTS AND RELEVANCE**

While CQADS has not been made privy to all details of the validation work conducted by CATSA on 2013 data (especially when it comes to exact accuracy figures), the CATSA Project Authority has let it be known informally that the predictions and QoS level curves which were generated by the WTIM were found to be quite in agreement with the actual data: it is CQADS’ understanding that the model is currently in use within Operations Reporting and Analysis at CATSA.

**PROJECT LOGISITCS**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Timeline** | |  |  |  |  | | --- | --- | --- | --- | | *Phase I* | Contractual Tasks | 14-Jun-14  31-Aug-14 | to | | Additional Tasks  (at CATSA’s request) | 01-Sep-14  30-Sep-14 | to | | *Phase II* | Contractual Tasks | 11-Oct-14  11-Nov-14 | To | |
| **Resources/Personnel** | Yiqiang Q. Zhao, Ph.D.  Professor, School of Mathematics and Statistics  Carleton University  Subject Matter Expert (Queueing Theory)  Patrick Boily, Ph.D.  Managing Consultant, CQADS  Project Lead / Senior Analyst  Wenzhe Ye  Consultant, CQADS  Analyst (Queueing Theory)  Jun Gao  Junior Consultant, CQADS  Analyst |
| **Total Effort Level** | |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | 580 hours (estimate) | | **Zhao** | **Boily** | **Ye** | **Gao** | | *Phase I* | Contractual Tasks | 60 | 100 | 120 | 50 | | Additional Tasks  (at CATSA’s request) |  | 100 |  |  | | *Phase II* | Contractual Tasks |  | 150 |  |  | |  | **Total:** | **60** | **350** | **120** | **50** | |
| **Dollar Value** | |  |  |  |  | | --- | --- | --- | --- | | *Phase I* | Contractual Tasks | $22,000.00 |  | | Additional Tasks  (at CATSA’s request) | $22,000.00 |  | | *Phase II* | Contractual Tasks | $10.000.00 |  | | **Total:** | | $54,000.00 | (+ HST) | |
| **CATSA Project Authority** | Maryam Haghighi, Ph.D.  Senior Advisor  Operations Reporting and Analysis  99 Bank Street  Ottawa, Ontario K1P 6B9  Canada  613 993-7094  [maryam.haghighi@catsa.gc.ca](mailto:maryam.haghighi@catsa.gc.ca) |

## Project Summary – Transport Canada

**Analysis of Fluidity Indicators and Seasonality Adjustments for Containers Transit Times in a Multi-Modal Supply Chain Networks**

by Patrick Boily, Yue Huang

Centre for Quantitative Analysis and Decision Support, Carleton University

**CLIENT ORGANIZATION**

Now, more than ever, Canadians need a safe and secure transportation system. Transport Canada (TC) is a government agency that is responsible for transportation systems, policies and programs. It promotes safe, secure, efficient and environmentally-responsible transportation within Canada and reports to Parliament and Canadians through the Minister of Transport.

As a result of ensuring a safe and secure transportation system, TC’s work protects people from accidents and exposure to dangerous goods, protects the environment from pollution, and contributes to a healthy population and economy. TC is also responsible for the safety and security of activities including: aircraft services, rail, road and marine safety, and transportation of dangerous goods.

**PROJECT INTENT, SCOPE, AND OBJECTIVES**

Supply chains play a crucial role in the transportation of goods from one part of the world to another. As the saying goes, “a given chain is only as strong as its weakest link” – in a multi-modal context, comparing the various transportation segments is far from an obvious endeavour.

TC is looking to produce an index to track container transit times in multi-modal chain networks. This index should depict the reliability and the variability of transit times but in such a way as to be able to compare performance between differing time periods. The seasonal variability of performance is relevant to supply chain monitoring and the ability to quantify and account for the severity of its impact on the data is thus of great interest.

The ultimate goal of this project was to compare quarterly and/or monthly performance data, irrespective of the transit season, in order to determine how well the network is performing, as it applies to the *Shanghai* → *Port Metro Vancouver/Prince Rupert* → *Toronto* corridors, and to produce a scoring methodology which could then be applied to other corridors.

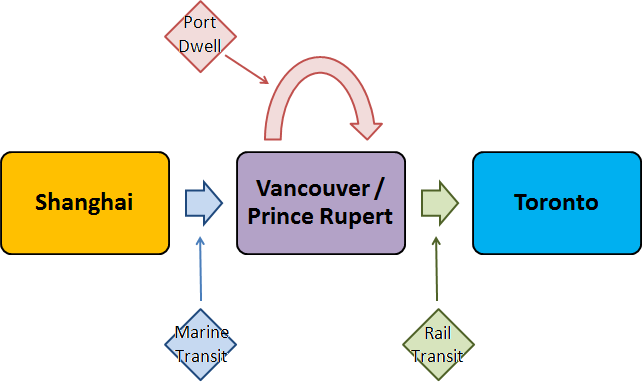
**METHODOLOGY**

In order to complete the assignment, CQADS used the following methodological steps:

1. *Review transportation literature*, in order to develop a scoring methodology to determine which is most relevant. The scoring methodology was applied to several proposed indicators that were developed for the *Shanghai* – *Port Metro Vancouver* – *Toronto* corridor.
2. *Review and explore available transit time data* to identify seasonality components, leading to adjusted data elements and to the elimination the variability component attributable to such trends.
3. *Testing various pre-existing reliability/variability indicators* against collected container transit time data, which identified promising leads.
4. *Development of the conceptual time-series model,* which established the logic and interaction between the proposed model indicators, and identified the essential data elements providing the best fit to the available data.
5. *Implementation of conceptual model* on a SAS platform, which allowed for the recognition that indicators which best reflected the performance of the chain in a given link were not necessarily the best choice for other links, and led to a new iteration of the prototype model.
6. *Final validation of the prototype model* using collected container transit time data, adjusted to reflect all the underlying trends that had been discovered.
7. *Documentation of the final model*: a technical report providing a detailed description of the model, as well as a number of useful scoring examples, was written and delivered to TC stakeholders. Quality assurance was insured by getting the report summarized and reviewed by a third party, external to the project.
8. *Knowledge transfer* was achieved through regular phone meetings and email exchanges detailing the project progress, and by getting the report reviewed and summarized by external parties.

**PROJECT SUMMARY**

The supply chain under investigation has Shanghai as the point of origin of shipments, with Toronto as the final destination; the containers enter the country either through Port Vancouver or Prince Rupert. Containers leave their point of origin by boat, arrive and dwell in either of the two ports before reaching their final destination by rail. The situation is illustrated in Figure 1 below.



**Figure 1** – The Shanghai → Vancouver/Prince Rupert → Toronto supply chain**.**

For each of the three segments (Marine Transit, Port Dwell, or Rail Transit), the data consists of the monthly empirical distribution of transit times from January 2010 to March 2013 (for Port) or April 2013 (for Marine and Rail), built from sub-samples (assumed to be randomly selected and fully representative) of all containers entering the appropriate segment.

Each segment’s performance was measured using Fluidity Indicators, which are computed using various statistics of the transit/dwelling time distributions for each of the supply chain' segments. The main indicators under consideration were:

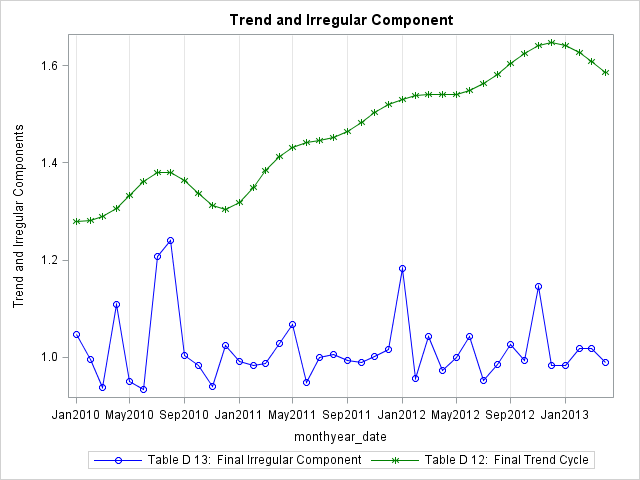
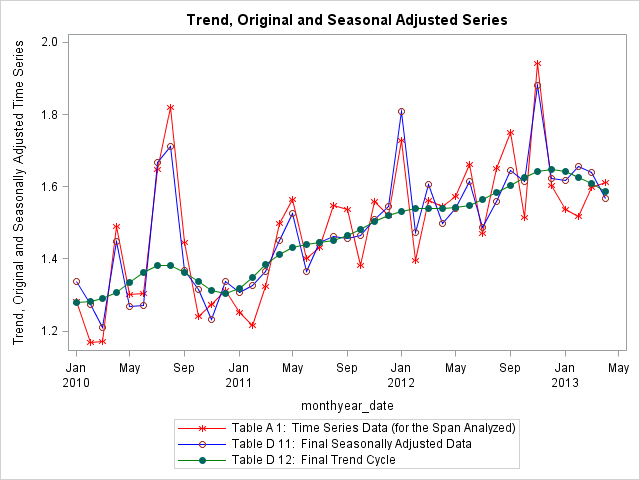
* the *Reliability Indicator* (RI) is the ratio of the 95th percentile to the 5th percentile of transit/dwelling times. A high RI indicates high volatility, whereas a low RI () indicates a reliable corridor;
* the *Buffer Index* (BI) is the ratio of the positive difference between the 95th percentile and the mean, to the mean. A small BI () indicates that the mean and the 95th percentile transit times are roughly the same, and so that there is only slight variability in the upper (longer) transit/dwelling times; a large BI indicates that the variability of the longer transit/dwelling times is high, and that outliers might be found in that domain;
* the *Coefficient of Variation* (CV) is the ratio of the standard deviation of transit/dwelling times to the mean transit time.

The time series of monthly indicators (which are derived from the monthly transit/dwelling time distributions in each segment) were then decomposed into their Trend, Seasonal (Seasonality, Trading-day, Moving-holiday), and Irregular components (see Figure 2, next page, for an example).

The Trend and Seasonal components provided the expected behaviour of the indicator time series; the Irregular components arose as a consequence of supply chain volatility.

Broadly-speaking, the decomposition involved three main steps:

1. the selection and application of a **seasonal decomposition model** (either additive or multiplicative), through *graphical inspection* (multiplicative if the size of seasonal peaks and troughs changes as the trend changes, additive otherwise) and/or *AICC comparison* (using the SAS procedure X12);

****

**Figure 2 –** Indicator time series for Marine RI (Shanghai → Vancouver) in red**.** The Trend component is shown in green, the Irregular one in blue.

1. the identification of, and adjustment (as required) for **calendar effects** such as **trading-day** **effects** (the effect of the monthly number of weekend days)or **moving-holiday effects** (due to Easter, for instance, in the Western world, or the Chinese New Year in the Pacific Rim), using *spectral plots*, *AICC tests* and *graphical inspection of diagnostic plots;*
2. the identification of, and adjustment (as required) for **trend level shifts** (abrupt but sustained changes in the underling level of the time series which usually have an identifiable cause, such as an increase in shipments due to an extra terminal having opened) and **outliers** (extreme values which fall outside the general trend pattern which can be caused either by an extreme random effect or an identifiable reason such as a short strike or a bad weather event), using *month-to-month percentage changes* and *residual patterns*.

[Time series decompositions, and hence any activity depending on them (such as forecasting, for instance), ultimately rely on the quality of the underlying data. In particular, there are a number of well-known data quality issues which affect the results of the analysis:

* the method of data collection may lead to unusual effects, especially if collection is made on a non-calendar basis or if there is a lag between activity and measurement;
* any change to the method or timing of data collection could lead to the false identification of trend or seasonal breaks;
* some series are sensitive to events such as extreme weather, strikes, wars, etc., which could cause breaks or outliers of large magnitude;
* at least 5 years’ worth of data are required to insure stability on future updates, and
* at least 10 years’ worth of data are required to insure that the adjustment of the first year is unlikely to be revised.

The last two of these did apply to the indicator time series, but were not deemed problematic in the long term as the data collection program from which the data was obtained was slated to continue indefinitely.

A larger issue, and one that it is more difficult to ignore, is that transit/dwelling times are only available for a sample of all containers going through the supply chain, and it is not at all obvious that this sample is randomly selected by the relevant authorities (and so may fail to be representative). Even if it was randomly selected, there is no guarantee that the sampling has remained (or will remain) the same over time.

Consequently, the analysis results were only ever as good as the quality of the data that went into the model. Since the aim of the project was solely to provide a methodology for time series decomposition – rather than to highlight particular irregular values on specific indicator time series – the issue is of lesser consequence at this stage. But this will have to be resolved one way or another by TC and its clients.]

There were no overarching results that apply to all indicator time series, on each segment (save for the lack of effect of the Chinese New Year, surprisingly enough): there were series with an Easter effect for a given indicator but not for another; series with a trading-day effect in a segment but not in another; series with outliers and series without, series with a trend level shift and series without.

[Another issue is that the reliability of a supply chain is a function of the total transit time from its origin to its destination. The importance of obtaining end-to-end data (i.e., of following a container from one end to another) has been recognized recently, and this data will be used in the analyses when enough of it becomes available. For the time being, however, the segmented data must somehow be joined together, one after the other, in order to provide approximate end-to-end data.

Once the seasonal adjustments are made on the segmented data (i.e. the Marine transit, Port dwelling, and Rail transit time data), we construct an **aggregate indicator** (or index) using these seasonally adjusted segmented indicators as a (necessarily poor) substitute for a direct end-to-end indicator, the underlying argument being that the total fluctuation for the supply chain should be the sum of the fluctuations from the segments since the supply chain as a whole is made up of the individual transportation modes.

Hence, the aggregate index is conceptualized as a weighted average of the indices of the component transportation modes:

where

* is the aggregate index for the entire supply chain at time ;
* is the component index for the specific transportation mode (or segment) at time (one of RI, BI or CV, say), and
* is the weight assigned to mode at time (the weights must be internally consistent from mode to mode in order for the weighted average to have meaning).

For a given supply chain, the average transit time in each mode is considered a good candidate for the weights since it functions as a reflection of the importance of the specific transportation mode to the entire supply chain, and since the average transit time of the entire supply chain is the sum of the average transit time of the individual component transportation modes. Had financial data been available, the **value-added** (the product of quantity and price) would also have been a good choice.

In the absence of financial data, however, we may suppose that the cost of a container spending a certain amount of time in a given mode is proportional to the amount of time spent in that mode (with the understanding that the proportionality constant may differ from mode to mode); as such, it makes sense to use where

* is the quantity of containers through mode at time , and
* is the average amount of time spent in mode at time .

With these assumptions, the aggregate index is eventually defined as

although it is important to note that there is no easy way to validate this formula without end-to-end data.]

After seasonality adjustments, it also became possible to compare the performance of (and hence to attempt to differentiate) the various indicators (RI, BI, CV) on a segment-by-segment basis:

* *Marine Transit* – all indicators show increasing trends in the Shanghai → Vancouver channel, and they all identify an outlier for MAY2011 in the Shanghai → Prince Rupert. In both channels, RI had a less volatile seasonal component, while BI had a less abnormally irregular component. It was thus not possible to cleanly rank RI and BI, but the adjusted data suggested that CV would be a poor selection as the indicator of choice.
* *Port Dwelling* – BI was seen to be the less volatile of all indicators, in both Vancouver and Prince Rupert.
* *Rail Transit* – RI was shown to be less volatile in the Prince Rupert → Toronto channel, but not enough data was available to come to a conclusion in the Vancouver → Toronto channel.

The importance of eventually collecting end-to-end data was made clear, as no clear-cut consensus for all segments emerged, apart from the unsuitability of CV as an indicator.

A number of supply chain scoring metrics (scaled scores, comparison scores) were also provided (pitting the adjusted expected data against the actual data in different ways), but before 5-years’ worth of data is available, it is a somewhat artificial endeavour to select the optimal one.

**ISSUES**

The short timeline allocation (see Project Logistics, below) was a consequence of a now-discontinued internship program, in which promising graduate students were hired as interns by CQADS to work on small projects and paid at a discounted rate in exchange for course credit. Consequently, the project dollar value (see Project Logistics, below) was about a fourth as large as it would have been under normal conditions.

Under these same normal conditions, the Centre would have negotiated a 3-month period over which to complete the project, rather than the agreed-upon 4 weeks. The total level of effort would not have changed.

Furthermore, some unforeseen data quality issues emerged (with respect to the Port Dwelling time in Vancouver) and as a result, the deadline was pushed back, upon mutual agreement.

**RESULTS AND RELEVANCE**

The suggested scoring methodology provided TC’s Economic Analysis and Research (EAR) group with a basis for implementing seasonality identification, and compensation methods. It is also known that the final report was circulated by EAR to a select group of academic and industry contacts. It eventually made its way into the hands of Prof. Ata Khan, of the Faculty of Engineering and Design at Carleton University, who was impressed with the work and as a result enlisted CQADS’s assistance for a study on the transportation of dangerous goods on behalf of the Nuclear Waste Management Organization.

**PROJECT LOGISITCS**

|  |  |
| --- | --- |
| **Timeline** | 10-May-13 to 12-Jun-13  (the original deadline of 07-Jun-13 was pushed back upon mutual agreement from both parties, given unexpected data issues) |
| **Resources/Personnel** | Patrick Boily, Ph.D.  Managing Consultant, CQADS  Project Lead / SME / Senior Analyst  Yue Huang  Consultant, CQADS  Analyst (Time Series Analysis and Forecasting) |
| **Total Effort Level** | |  |  |  | | --- | --- | --- | | 250 hours (estimate) | **Boily** | **Huang** | | **Total:** | 125 | 125 | |
| **Dollar Value** | $4,424.73 (+ HST) |
| **TC Project Authority** | Alexander Gregory  Economic Analyst  Economic Analysis and Research  4900 Young Street  North York, Ontario M2N 6A5  Canada  416 973-2444  alexander.gregory@[tc.gc.ca](mailto:maryam.haghighi@catsa.gc.ca) |

## Project Summary III – Forbes Laboratory

**Discrimination and Non-Discrimination in Parasites: A Simulation Created for Analysis of Evolutionarily Stable States and Population Dynamics**

by Jennifer Schellinck

Sysabee

**CLIENT ORGANIZATION**

Mark Forbes is a researcher at Carleton University and a world-recognized expert in parasitology, with an extensive research program investigating parasite-host behaviours, population dynamics and evolution of parasites and hosts, considering both theoretical and empirical questions in this context.

His research laboratory at Carleton University conducts research into these topics using various methodologies, including experimental research, theoretical research, field studies, mathematical models and computer simulations.

**PROJECT INTENT, SCOPE, AND OBJECTIVES**

The intent of this project was to understand which environmental, parasite and host factors – including properties and behaviours – might influence the evolution of particular parasite traits and abilities, specifically the ability or non-ability to differentiate amongst and choose between different types of hosts, the selection of which would then have differing influence on the reproductive success of individuals within the population of interest.

Another area of interest was the investigation of how these factors influence the dynamics at a population level (i.e. which parameter settings resulted in the appearance of certain attractor states within the simulation).

The scope of the project was intended to encompass, and be generalizable across, a wide variety of host and parasite species. The results of the project were broadly applicable to a number of parasitology and epidemiological questions relating to parasite spread in humans and other species. The study investigated multiple possible evolutionary patterns relating to the evolution of host discrimination or host non-discrimination within parasite populations over time, using a multi-agent simulation methodology.

The simulation modeled system behaviour at an individual, genetic level, with evolution occurring within the population as a result of reproduction and genetic cross-over of individuals. Population level dynamics were emergent from these individual level behaviours. The project investigated numerous scenarios, including haploid vs diploid genetics, invasion of one parasite sub-species into a population of a second parasite species and environments with varying levels of host presence.

The client’s objective was to investigate how individual behaviour and genetics influence population dynamics and evolution of parasite traits over time. In order to do this, the client needed to track the behaviours and properties of specific parasites with specific traits and then determine how these behaviours and properties, in combination with environmental factors (e.g. availability of host types within the environment) influenced reproductive success of individual parasites, as well as how these individual behaviours then generated emergent population level properties, dynamics and states. The client also needed to be able to generalize these findings to more than one type of parasite and host species.

**METHODOLOGY**

Broadly speaking, the project used the following methodologies:

* multi-agent stochastic simulations;
* data and statistical analysis via customized computer programming (using Perl), and
* data visualization of stochastic simulation behaviours and large stochastic datasets.

With respect to modeling, the project used a methodology developed by Schellinck and Webster (2010, 2013) in order to support quantitatively accurate and scientifically rigorous simulations. In accordance with this framework, Dr. Schellinck used the following methodological steps:

1. Definition of model purpose
2. Definition of model scope
3. Target phenomenon conceptualization and ontology

4a. Definition of conceptual level model

4b. Definition of model-real system analogies

5. Model implementation

Schellinck, J., Webster, R. (2013) “Cognitive models: Understanding their critical role as explanatory and predictive hypothesis generators in cognition research”. Poster presented at ICCM 2013 The 12th International Conference on Cognitive Modelling, Carleton University, Ottawa, Canada.

Schellinck, J., Webster, R. (2010). "The Scientific Power of Good Models: Unifying Hypothesis Discovery and Hypothesis Testing". Presented at Models and Simulations 4, University of Toronto, Toronto Canada.

**PROJECT SUMMARY**

The primary goal of this project was to shed new light on a poorly understood aspect of parasite behaviour and evolution- specifically, under what conditions would parasite evolution favour specialization to only one host species and, conversely under what conditions would parasites evolve the ability to exploit multiple host species.

The exploration of this question required the tracking of a large number of parasite traits and behaviours, at an individual level, and in a tightly controlled environment. These specific needs led to issues with typical biological research methods, such as laboratory experiments and field research, which would not allow for the environmental control and precision data gathering required to generate an understanding of parasite-host dynamics that could meet the client objectives.

Overcoming these issues necessitated the use of interdisciplinary methodologies involving specialized computer science and cognitive science techniques (multi-agent modeling) along with systems dynamics analysis techniques in order to understand the resulting behaviour of the simulation. Dr. Schellinck’s contribution was to provide an individual who was experienced in designing, implementing and analyzing simulations that would allow the client to meet the stated objectives.

With respect to complexity of the simulation, the resulting simulation, which was multi-agent in nature, involved modeling the individual behaviours of up to 1600 parasite agents over up to 2000 timesteps for each set of parameters being investigated. The design of the simulation was multi-factorial in nature, investigating numerous factors and numerous parameter settings for each factor.

Within agents, behaviour was based on genetics level encoding, which was itself determined for each individual agent by implementing a biologically realistic reproduction algorithm in order to accurately model biological evolution. Over the course of the research project simulation parameter exploration generated over a million data points which were then analysed statistically in order to come to generalizable conclusions about host-parasite evolution of discrimination and non-discrimination.

**DIFFICULTIES AND ADDITIONAL TASKS**

The project experienced, and successfully overcame, many of the common difficulties and challenges associated with an interdisciplinary modeling project requiring the creation of a simulation that accurately represented a complex and multi-faceted real world problem. Particular challenges addressed during the project included:

* Establishing successful communication between modeling expert and subject matter expert- in this case the simulation and system dynamics expert (Dr. Schellinck) and the parasitology expert (Dr. Forbes). This was made possible by communicating in advance the intended methodology which would be followed during simulation creation, as well as a willingness on the part of the subject matter and simulation expert to clearly define key concepts and theoretical constructs in such a way that they could be accurately operationalized in a simulation context.
* Ensuring that the model was accurate to a level that resulted in predictions that were also sufficiently accurate- this was managed by ensuring that model behaviour and parameter settings at the individual level were supported and established based on existing published research on parasite physiology and behaviours.
* Ensuring that the aspects of the situation of interest that influenced the behaviour of the model were sufficiently included in the model – this required a fairly extensive transfer of knowledge between the subject expert and modeling expert.
* Understanding how model results can be, and should appropriately be, reapplied to the situation of interest. Again, this required a careful discussion of which results had relevance to the parasitology community, as well as how systems dynamics results could be successfully communicated to this community.
* Understanding the implications when model results are not consistent with either known or expected situations. This required a thorough knowledge of the relevant research literature, and an ability to extrapolate this to the sometimes novel behaviours exhibited by the simulation.

In general, these challenges were anticipated and successfully managed through application of the modeling framework discussed in the methodology section, above.

**RESULTS AND RELEVANCE**

By the completion of the project a number of relevant conclusions about the evolution of discrimination and non-discrimination within parasite populations, along with the factors that lead to the evolution of these traits, were drawn from analysis of simulation behaviours. Specific conclusions, as presented at Genomes to/aux Binomes 2014, were:

* Except in extreme circumstances, both non-discrimination and discrimination are relatively successful strategies in the majority of environments.
* The continuing presence of non-discriminators in environments that favour discrimination means that parasites will remain adaptable to new circumstances.
* In most circumstances we would not expect to see abrupt shifts either away from, or towards, discrimination or non-discrimination.

More specific quantitative conclusions from the simulation were:

* Any shifts that do occur will occur over > 100 generations (which may represent a time span of anywhere from weeks to years, depending on species).
* Most scenarios favour discrimination, but in these scenarios non-discriminators can remain present for multiple generations (500+ generations).
* When the cost of discrimination is high, discrimination will quickly die out when non-discrimination is introduced (within < 100 generations).
* When one host can’t be exploited at all, there is a relatively high selection pressure in favour of discrimination.

The broader relevance of this research relates to an increased understanding of both the likelihood of, and environmental conditions required for, a particular parasite species to evolve either from a discriminating species into a non-discriminating species, or from a non-discriminating species into a discriminating species. In the case of human host species, either of these outcomes may lead to several epidemiologically relevant implications under particular environmental conditions, including:

* An increase in the level at which humans become parasitized
* An increase in the accidental parasitization of non-host species, including humans
* Novel spread of pathogens between species – i.e. humans are infected with novel pathogens when parasites evolve into a non-discriminating species
* Exploitation of entirely new hosts via evolution into generalists – i.e. parasites which previously did not use humans as hosts may start using humans as hosts
* heavier parasitization of the preferred host species, or even of particular individuals within that species
* aggregation of strains leading to speciation- i.e. the evolution of new parasite species

As a result of the potentially problematic implications of such scenarios, it is important to increase our level of understanding of the conditions under which these particular outcomes may be more or less likely to occur, via simulations research such as that carried out for this project.

**PROJECT LOGISITCS**

|  |  |
| --- | --- |
| **Timeline** | Phase 1: 08-May-2007 to 21-Dec-2008  Phase 2: 03-May-2013 to 26-May-2014 |
| **Resources/Personnel** | Jennifer Schellinck, Ph.D.  Principal, Sysabee  Project Lead / SME / Senior Analyst |
| **Total Effort Level** | 190 hours (estimate) |
| **Dollar Value** | $15,000.00 |
| **TC Project Authority** | Mark Forbes  Associate Vice-President (Research)  Tory 503 Lab: 234  Nesbitt Building  1125 Colonel By Drive  Ottawa, ON, Canada, K1S 5B6  613 520-3570  [mark\_forbes@carleton.ca](mailto:mark_forbes@carleton.ca) |

3. Report Sample

The sample is extracted from the CQADS report *Wait Time Impact Model at Pre-Board Screening Checkpoints for Canadian Airports*, by Yiqiang Zhao, Patrick Boily, and Wenzhe Ye.

The final version of this report was presented to CATSA in December 2013; it describes the Wait Time Impact Model (WTIM) that was developed by CQADS (see Project Summary I for a list of project objectives).

The following 5-page extract contains passages from §1.2, §2.1, §2.2, §3(.0), §3.1, §3.2, §5.3 and §8(.0). Passages which have been skipped are indicated in the text by the symbol “[…]”. Non-contiguous passages were selected in order to showcase the point-rated criteria 3.1 and 3.2 (RFSO Schedule “C”, item 2.5).

The colour scheme has been modified from red accents to blue accents in order not to create confusion between the section numbering of the extract and the section numbering of the current bid proposal.

The extract cites the following references:

1. Newell, G.F. [1971], *Applications of Queuing Theory*, Chapman and Hall.
2. Ross, S.M. [2010], *Introduction to Probability Models*, 10th ed., Academic Press.

**Note:** The 35-page report was produced with LATEX, a typesetting system used to prepare documents of a mathematical nature. Due to the RFSO’s formatting requirements, the extracted sample is presented in the Arial typeface, except for the equations, which are presented in a Cambria typeface in order to preserve legibility. The paragraphs have been lightly re-arranged to meet those same formatting requirement, and some typographical errors have been corrected.

Extract from the report

**Wait Time Impact Model at Pre-Board Screening Checkpoints for Canadian Airports**

by Yiqiang Zhao, Patrick Boily, Wenzhe Ye

[…]

**1.2 Model Outline**

The model establishes a relationship between the arrival rates, the service rates, the number of servers and the service levels. Basic concepts, process descriptions, and limitations are provided in §2.

The WTIM is best described via the flow chart of Figure 1 on the next page (the various concepts will be defined as they arise in the corresponding section):

1. computation of the arrival rates from the raw data (§3.2);
2. computation of the distribution of the number of servers from the checkpoint utilization reports (§3.3);
3. computation of the waiting time distribution from the waiting time report (§3.4);
4. computation of the QoS levels from the waiting time report (§3.4);
5. computation of the estimated QoS levels under the assumption (§3.5);
6. validation of the assumption based on a comparison of and (§3.6);
7. computation of the estimated service rates under the assumption (§3.5);
8. computation of the seasonal checkpoint regression parameters under the combined and *Regression* assumptions (§4.1);
9. computation of the estimated QoS levels under the combined and *Regression* assumptions (§4.2);
10. validation of the combined and *Regression* assumptions based on a comparison of , and (§4.3);
11. prediction of the number of servers under the combined and *Regression* assumptions (§5);
12. validation of the combined and *Regression* assumptions based on a comparison of and (§5.3);
13. computation of the checkpoint departure parameters under the combined , *Regression* and *Departure* assumptions (§5.3);
14. computation of the estimated QoS levels for various projected arrival growth rates under the combined , *Regression* and *Departure* assumptions (§6.2);
15. prediction of the number of servers for various projected arrival growth rates under the combined , *Regression* and *Departure* assumptions (§6);
16. final validation of the combined , *Regression* and *Departure* assumptions based on a comparison of and with empirical data (§6.3).

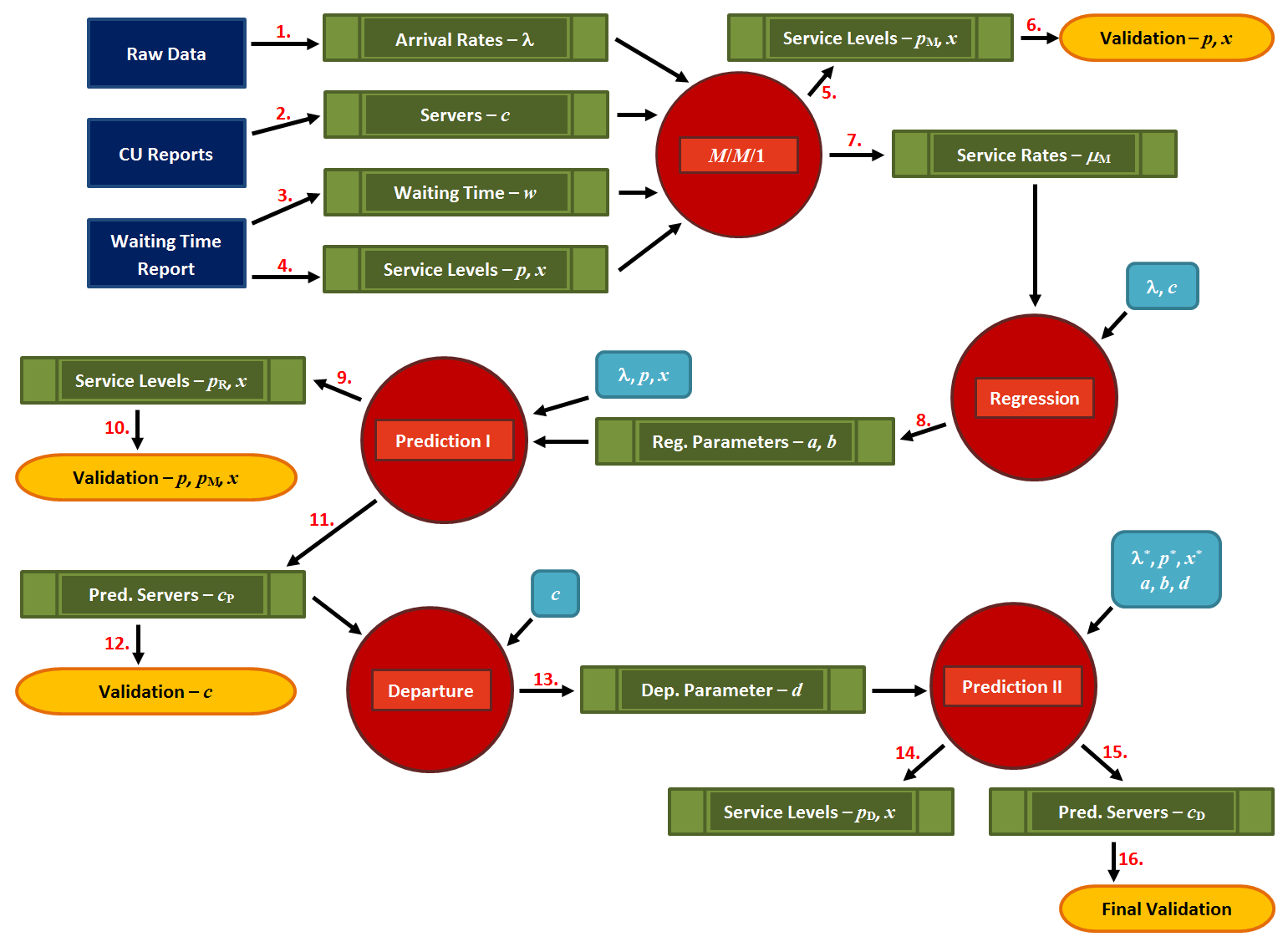
In order to illustrate the WTIM process, the details are worked out on a step-by-step basis for the Domestic/International Checkpoint at the Edmonton International Airport (YEG), based on 2012 data. The results are shown at the end of each section. A summary of results for all checkpoints is also provided, as well as recommendations and suggested next steps.

[…]

**2.1 Definitions**

The various mathematical concepts to which the report will refer are described below:

* An **queueing model** describes a system where arrivals form a single queue and are governed by a Poisson process (the first ), units arriving are processed by servers and service times are exponentially distributed (the second ).
* A **Poisson process** is a stochastic process where the time between any two consecutive event has an exponential distribution with parameter .



**Figure 1** – **WTIM flow.** The dark blue rectangles are CATSA-provided data inputs; the green boxes indicate computed and derived values; the red circles are conceptual nodes; the light blue boxes represent carry-over values, and the orange cells are validation steps.

* The **arrival rate** is the rate at which passengers arrive for PBS (i.e. passengers per minute), the **service rate** is the processing rate at a screening line (i.e. maximal potential throughput), the **number of servers** is the number of screening lines and the **service level** is the percentage of people waiting less than a given number of minutes at a checkpoint.

**2.2 Description of the PBS Process**

At each checkpoint, the PBS process is structurally similar: passengers arriving at the beginning of the main queue may have their boarding passes scanned at the position, but they are always scanned at the position (see Figure 2 on the next page).

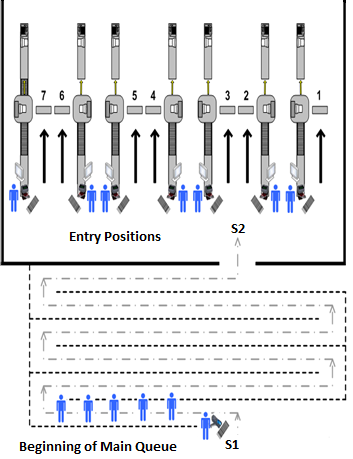
[…]

**3. Queueing Model**

One of the difficulties for the situation under consideration is that the number of servers varies with time, according to different factors: there are times when all servers are busy, others when a number of open servers are idle, and the number of open servers changes according to some vacation policy which it is difficult to model. This is problematic when using an model as service rate estimates depend, amongst other things, on the number of open servers.

It is possible to circumvent this issue altogether, without invoking Vacation Models, by noticing that an queueing system may be viewed as an queueing system where the servers are hidden behind a generalized server (see Figure 3, on the next page). Under that interpretation, the service rates can be estimated independently of the number of servers. Furthermore, not only do results still hold for (simply by setting in the appropriate theorems), but the quantities to be computed tend to be simpler in this case.

While this conceptual simplification has removed some of the difficulties associated to server vacation, there remains another problem: the theory of systems, alone, is not sufficient to recover (and later predict) the actual (and hidden) number of servers for the checkpoint. This situation can be addressed by finding another way to link the arrival rates, the estimated service rates and the number of servers (see §4.1 for more details).

**Figure 2** – **Schematics of pre-board screening (PBS).** Passengers enter the main queue, where their boarding pass may be screened at . Once they reach the end of the main queue, their boarding pass is screened at and they are sent to one of the active lines for processing (image provided by CATSA). In practice, it may happen that only the reading is available.

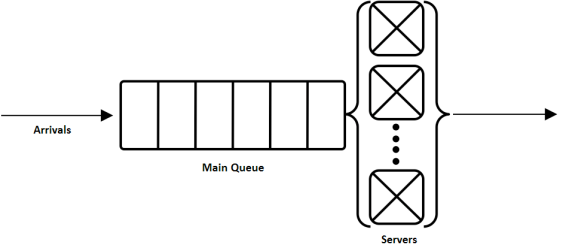
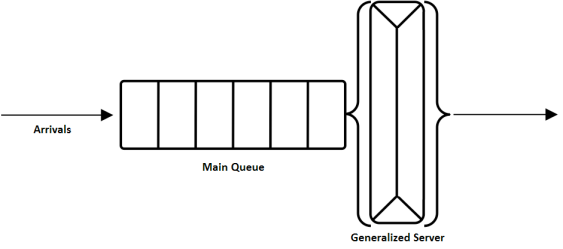
**3.1 Clustering**

In order to better predict the average behaviour of a system and its possible outcomes, a wide range of typical patterns must be considered. When analyzing the behaviour of queues, it may become necessary to group the data into meaningful clusters exhibiting similar properties (for example, properties that can be characterized by the same Poisson process).

This approach allows for proper estimation of queuing model parameters (arrival rates, processing rates, etc.), which in turn yields the most reliable results. The selection of the appropriate cluster size relies on finding a balancing point between two extremes:

* In order to properly define the stochastic process, a minimum amount of data with similar properties is required. If clustering is not performed (i.e., if the clusters are too large), the data may present different characteristics which cannot be represented by a single Poisson process.
* On the other hand, if the clusters are too small, they may not contain enough data to capture the underlying properties. More importantly, clusters that cover too short a period are unlikely to exhibit the statistical behaviour of the process.

A preliminary analysis of the model’s accuracy was assessed based on the following criteria: *Checkpoint*, *Weekly patterns* (day of week vs weekday/weekend), *Seasonal patterns* (season vs month) and *Daily patterns* (2-hour period vs 4-hour period). The cluster combination that produced the most encouraging queueing results when compared against actual reports was: checkpoint, weekday/weekend, season, 4 hour-period. Clustering also plays a role in the Regression stage of the model (see §4 for details), but the optimal regression cluster combination need not be the same as the queueing cluster combination.

  **Figure 3 – Queuing systems.** Conceptual visualization of an queue (on the left)as an queue (on the right); the servers can be considered as 1 generalized server. d

**3.2 Computing the Average Arrival Rate**

Since not all boarding passes are scanned at , the Wait Time report ( data) cannot be used to derive the cluster arrival rates. The line-up (main queue) is a birth-death process (i.e. a reversible one-dimensional Markov chain). In particular, the forward chain and its reverse are stochastically identical and the arrival epochs of the reversed chain are the departure epochs of the forward chain. We can then use Burke’s Theorem for queues.

**Theorem** **1** (Burke’s Theorem, [1]) *Consider an queue in the steady state with arrivals modeled by a homogeneous Poisson process with rate parameter . Then the departure process is also a homogeneous Poisson process with rate parameter .*

This does not rule out the possibility that, at a particular time, the arrivals at could be greater than the departures at , due to the inherent randomness of Poisson processes. But all arrivals will eventually leave at and thus the fluctuations at follow the same statistical property governing arrivals to the queue. Therefore, the arrival rates can be estimated by using data readings at within a given cluster.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |

Table 1 – YEG DI 2012 totals. Number of hours, count of arrivals and average arrival rates, per cluster, per quarter (1st – teal, 2nd – green, 3rd – yellow, 4th – red).

It remains only to show that arrivals follow a homogeneous Poisson process in each cluster (this is a common hypothesis). To do so, one must show, assuming that the number of arrivals in the cluster by time is denoted by , that (see [3, 4] for details)

1. is a counting process with independent and stationary increments, and
2. the number of arrivals in any time interval of length is Poisson-distributed with mean , i.e. for all

The first assumption is satisfied with the introduction of clusters. The third assumption holds if the **inter-arrival times** (the times between consecutive events) are independent and identically distributed (i.i.d.) exponential random variables with the same rate : analysis of in the raw data suggests that this is the case.

The total counts of arrivals for each cluster at YEG’s D/I checkpoint based on 2012 data are shown in Table 1 (above). Note that the arrival rate is simply calculated by dividing the count in each cluster by the number of minutes in each cluster, independently of the open status of the checkpoint during the period spanned by the cluster. A low arrival rate may thus indicate either that checkpoint traffic was low or intermittent for the cluster, or that it was closed for some or all of the period that it spans.

[…]

**5.3 Validating the Combined Model Using Departure at the Checkpoint Level**

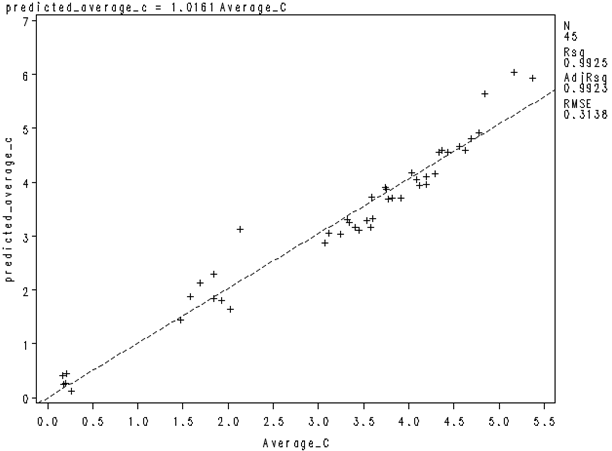
The relative accuracy of the formula

used to estimate the average number of servers required to reach the QoS level at a checkpoint with regression parameters and average arrival rate , suggests another method to validate the combined model.

For any given checkpoint, the plot of against the actual strongly suggests that the variables are linked according to , for some .

Linear regressions once again determine the optimal for each checkpoint. The **departure parameter** , then, serves as a measure of the predictive model’s departure from reality. If (i.e. if ), then the assumptions that go into the combined model are justified *a postiori*, in the context of predicting the average number of active servers. The modified predictions for a checkpoint where is large or close to 0 (i.e is a poor approximation of ) may still end up being accurate (i.e ), but in that case a careful analysis should be undertaken to understand whether any anomalous activity is in play.

The regression of against for YEG’s D/I checkpoint (based on 2012 data) is shown in Figure 6 (on the next page). The departure parameter for this checkpoint is , which is reflected by the tight linear fit of the two variables. As such, the original prediction requires only a very slight modification.

**Figure 6 – Departure Regression.** Regression of the predicted average number of servers against the actual number of servers.

[…]

**8. Recommendations and Final Comments**

Perhaps the foremost conclusion is that the model on its own provides the best QoS levels estimates, while the best estimates for the average number of active servers are provided by the Departure model.

This discrepancy may be partly explained by the fact that, in any modeling endeavour, some loss of information is inevitable due to the necessity of making simplification assumptions. Below is a list of possible issues which could affect the WTIM’s accuracy:

1. The underlying arrival processes are roughly Poisson, and the wait time distributions are roughly conditionally exponential for each cluster; depending on the distance between the theoretical process and the empirical data, the assumption may be inappropriate.
2. The wait time distribution may be seriously biased as not every boarding pass has been scanned at , and there is no easy way to verify how representative the subset of those for which wait time data is available actually is.
3. The server vacation policy is unknown, and may not be uniformly adhered to (if one even exists).
4. The actual number of active servers is only crudely approximated by the maximum number of active lines within a 15 minute block.
5. The service rate seems to depend on factors other than the number of active servers and the arrival rate, leading to wildly different outputs for similar inputs and contributing to the lessened accuracy of the regression model when estimating QoS levels.
6. Different checkpoints might require different optimal clustering strategies.

It might be possible to minimize some of that information loss simply by selecting a slightly more sophisticated regression functional form linking the average arrival rate per line and the average service rate per line. Preliminary analysis suggests that the choice

may provide better QoS results. Further analysis is needed in that regard, as it is clear that other factors need to be included in order to get the best possible fit and to minimize the number of clusters which become unstable as a result.

Finally, it is conceivable that while adding more historical data to the model could have a useful effect, going too far back into the past may bias the results if policy changes have led to characteristically distinct underlying data over the years. It seems clear that at least one year’s worth of data is needed, but, as the datasets only contained trustworthy data for the year 2012, it is still too early to get a definitive answer on this topic.

[…]

(End of extract)

4. Quality Assurance and Quality Control

CQADS acknowledges its responsibility to ensure the quality of its analyses and credibility of the interpretation it present to its clients (when applicable).  Accordingly, the Centre commits to providing rigorous and high-quality analysis and reporting, and to seek feedback from the CATSA Project Authority (PA) and stakeholders to ensure that their needs are being met.

The quality assurance process includes peer and expert review (subject to security and non-disclosure considerations), so that every deliverable is reviewed by at least one SME. Furthermore, the proposed team members have been selected for their expertise and proven track records in producing high quality research and deliverables.  Project work is not conducted in isolation and every product represents the combined efforts, expertise and consensus of a number of team members.

Quality control procedures allow CQADS to address any discrepancies or omissions as well as prevent delays: any CATSA concerns are immediately heeded, and steps to rectify these are enacted.  In spite of the Centre’s due diligence, it is conceivable that delays will occur: the PA will be notified as soon as the potential for delay is identified.  Furthermore, should additional resources be required, other analysts with comparable expertise can be called upon (once they have been vetted by CATSA).

Progress reporting to CATSA will include:

* *Regular touch-base meetings*, in the form of brief phone conversations between the CQADS Project Lead and the PA to communicate progress and troubleshoot issues, and
* *Bi-weekly progress reports,*  to be sent by the CQADS Project Lead to the PA, documenting, as a minimum, the following information:

1. activities that took place during the last two weeks;
2. activities that will take place during the coming two weeks;
3. a description of any problems encountered which are likely to require attention;
4. any recommendations relating to the conduct of the work; and
5. upcoming project dates/milestones.

* *Occasional longer sessions* involving the Project Team and CATSA stakeholder to demonstrate software, to transfer key knowledge and/or information (in both directions), to discuss internal and external validation strategies, and to present milestone results.

It has been CQADS’s experience that open, transparent, and frequent exchanges between all involved parties help to avoid time-consuming mistakes in the first place (**quality assurance**), and to detect some of the less obvious data and/or analytical flaws (**quality control**) – in short, they help identify, recognize and successfully counter anticipated risks.

In more practical terms, possible project products include research products (e.g. data and research studies), models (e.g. simulations, mathematical models, analytical and conceptual models) and analyses products (e.g. statistical analyses, operational analyses, sensitivity analyses). For each of these products, typical risks and constraints, and the resulting quality assurance and control measures to address these risks and constraints are outlined below:

**Research products (data and research studies)**

*1. Examples of typical risks and constraints*

* Required data is unavailable
* Existing data is problematic in some way (e.g. incorrectly collected, not ‘clean’, biased)
* Relevant existing research or information is not incorporated into the research study being carried out
* Relevant existing research is not interpreted correctly or interpreted in a biased or non-objective manner when being incorporated into a research study

*2. Quality Assurance Procedures*

* Existing data will be assessed for quality prior to analysis
* Any necessary data collection will be carried out based on plans devised to maximize data quality and ensure adequate data quantity and sampling accuracy
* Literature reviews will be conducted to determine the industry/academic consensus, and whether it applies to the situation at hand.
* Multiple individuals will be involved in the implementation of research studies, with the findings of individuals being compared and contrasted to ensure comprehensive and objective interpretation of results

*3. Quality Control Procedures*

* Scope and quality of data will be reviewed with CATSA in the case of existing data, and verified by multiple project members in the case of newly collected data. In the case of existing data obtained from the client, a data quality interview will be carried out to establish a baseline for data quality and completeness.
* Review of resulting research study by a non-author to validate the completeness of the research report as well as the objective nature of the results obtained.

**Models (simulations, mathematical models, analytical and conceptual models)**

*1. Examples of typical risks and constraints*

* Model does not accurately reflect situation of interest
* Model is too abstract to accurately or precisely answer question of interest
* Model is too granular to be constructed properly given existing amount and granularity of data,
* Incorrect analogies are drawn between the model and situation being modelled
* Model results misinterpreted or misapplied.
* Model variable values or simulation parameters are set on an ad hoc basis rather than being based on empirical data
* Functional nature of the model (e.g. analytic vs simulation) does not allow for adequate exploration of scenario dynamics
* Computational constraints make it difficult to thoroughly explore identified relevant scenarios, or difficult to explore all relevant scenarios.

*2. Quality Assurance Procedures*

* Literature review (redux)
* Models will be constructed using methodologies designed to ensure that accurate analogies are drawn between the model and situation.
* The required output of the model, and level of detail and accuracy will be verified prior to the commencement of model construction to ensure appropriate granularity
* Variable values and parameters will be validated based on existing empirical research findings or analysis of collected data
* Prior to construction of any models, an assessment will be made as to the required purpose of the model and it will be verified that the chosen model type has adequate functionality to meet this purpose.
* Model construction will take into account likely computational requirements of the resulting model and existing computational resources, in order to enable appropriate scenario exploration upon model completion.

*3. Quality Control Procedures*

* Model construction and accuracy will be reviewed and validated by situational experts.
* Model and simulation scenarios will be reviewed by additional team members to confirm that all parameter settings are realistic and based on existing research
* Upon completion of the analytic model, its accuracy will be verified by situation experts. The type of scenarios that can be accurately explored by the analytic model will be reviewed and discussed to confirm that it has adequate functionality

**Analyses products**

*1. Examples of typical risks and constraints*

* Incorrect assumptions lead to the application of incorrect or inappropriate statistical tests
* Incomplete knowledge or misunderstandings of relevant components of the system being analyzed lead to incorrect assumptions
* Some relevant variables are overlooked during the sensitivity analysis, resulting in an incorrect assessment of the sensitivity of the situation to particular conditions

*2. Quality Assurance Procedures*

* Literature review (part III)
* Possible statistical analysis approaches will be identified based on functional requirements. Statistical assumptions involved will be explicitly identified and confirmation that the situation being analyzed meets these assumptions will be obtained prior to carry out the statistical test.
* Scope of analyses will be identified prior to analyses. Confirmation of adequate data collection will be confirmed with situational experts prior to carrying out analyses
* A careful assessment of all possible variables will be reviewed and additional variables not involved in the sensitivity analysis will be evaluated with respect to their feasibility and appropriateness for inclusion in or exclusion from the sensitivity analysis

*3. Quality Control Procedures*

* Statistical tests and outputs will be reviewed by additional members of the project team to confirm accuracy of tests and correct choice of tests.
* Upon completion of the operational analyses, accuracy will be verified by situation experts.
* ‘Sanity checks’ will be carried out by both team members and situational experts to determine if results of the sensitivity analysis is consistent with the current understanding of the system being analyzed. In the event of unexpected results the sensitivity analysis will be reviewed.

5. Understanding of Requirement

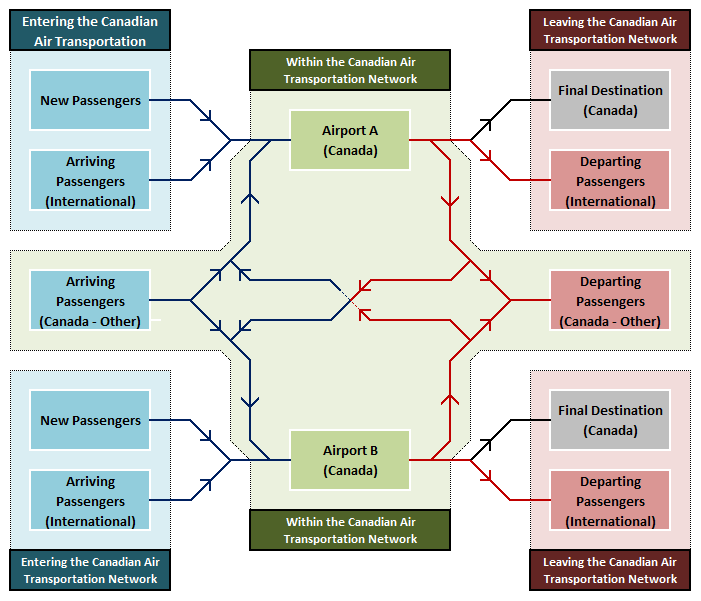
Broadly speaking, CATSA’s mandate is to ensure civil aviation security by carrying out a number of passenger screening services, including Pre-board Screening (PBS), Hold Baggage Screening (HBS), Non-Passenger Screening (NPS) and Restricted Area Identity Cards (RAIC) (sub-contracted to third party security agency firms)..

This security requirement is balanced by the need to process screenings consistently efficiently, and cost-effectively, without grinding the flow of air travellers through Canadian airports to a halt, nor putting them in jeopardy.

To achieve their mandate, CATSA requires ongoing analytics support, which includes

* operational analysis of the security environment;
* construction of models and simulations to understand current operations and explore new security options,
* the statistical analysis of data collected on factors influencing operations;
* the maintenance and upgrading of existing models created to predict security operation activity.

But what does this traffic look like? Generically, the flow could be described by a graph not unlike the one shown in Figure 1 (below).

**Figure 1** **–** **Schematics of passenger traffic flow in the Canadian Air Transportation Network.** 2 nodes (Airports A, B) are highlighted; blue arrows indicate traffic going *into* a node, red arrows traffic going *out of* a node, black arrows indicate traffic leaving the Network.

The flow through each Airport might be described as in Figure 2 (next page).

In order to facilitate CATSA’s ability to make strategic decisions based on this picture of the Canadian Air Traffic Network, several theoretical fields and analytical processes must be brought together in a practical manner, providing insight into operational issues, the ability to ask the right questions, and mostly, useful answers.

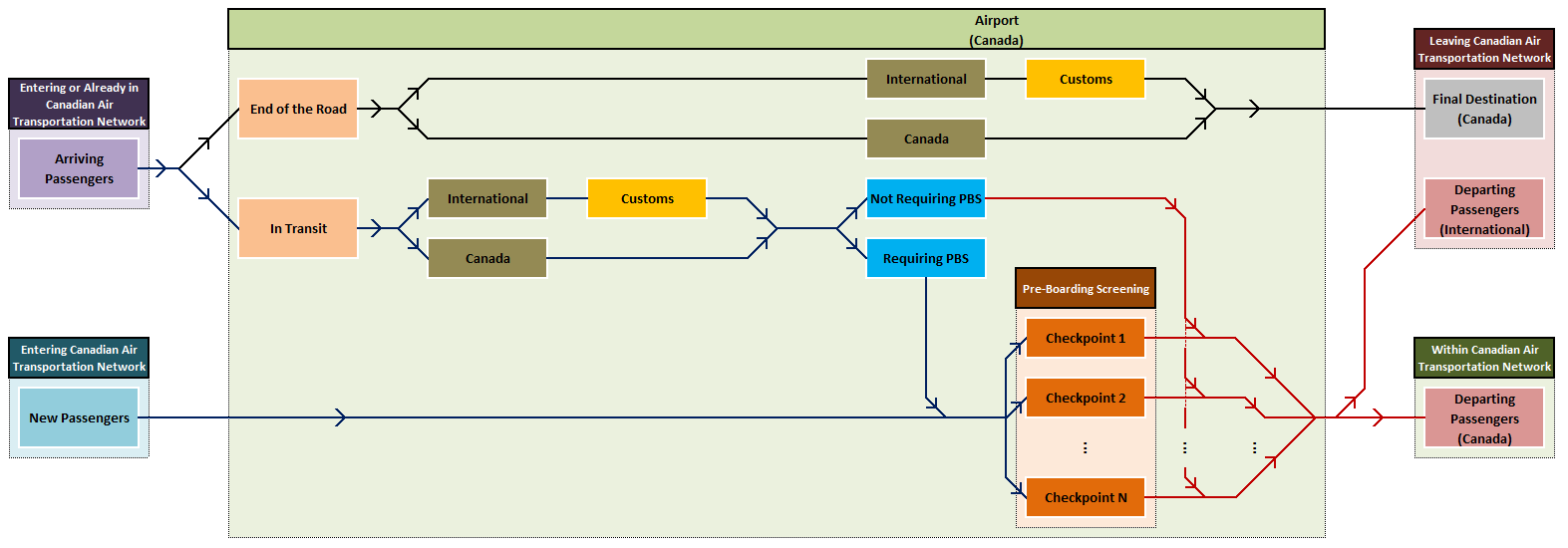


Figure 2 – Schematics of the traffic flow through a node (airport). Arrows are defined as in Figure 1.

Suppose, for instance, that CATSA would like to find out more about the potential effects of a given policy on the network (the specific details of the policy are not essential at this point).

It would be impractical (not to mention possibly hazardous) to gauge the effect by implementing the policy and analyzing its impact on the actual traffic flow. An alternative approach might consist of building an integrated simulation model which could test various policy scenarios and help determine which of them should be considered for the next implementation/testing phase at CATSA’s checkpoints.

Such a **simulation** approach might require the integration of the following modules:

1. A method to **forecast** the traffic flow through each checkpoint (or of the number of flights into and from an airport) over time, based on a number of external factors (such as the state of the economy, population growth, etc.);
2. A **predictive model** to determine origin/destination pairs for travellers, based among other things, on the forecast provided by Step 1;
3. A **queueing model** which predicts either the wait time distribution at each checkpoint, based on the forecast and predictions of Steps 1–2., or the number of servers required to meet a benchmark Quality of Service (QoS) level for the waiting time, again based on the forecast and predictions of Steps 1–2;
4. An **optimization** routine which allocates staff to the various processing servers in order to meet required QoS thresholds based on the forecast, predictions, and wait time distributions of Steps 1–3 and available budget, say;
5. A method to evaluate the network’s performance, or to validatethe network model, by comparing **statistical properties** of the model outcome (Steps 1–4) with those of available empirical data, when available;
6. A **statistical,** **evidence-based** comparison/evaluation of the various policy scenario outcomes (Steps 1–5).

On the strength of a previous contract with CATSA, CQADS is aware that some work has already started on queueing models to describe the wait times at PSB checkpoints.

Operational analysis could also provide insight into security procedures and processes, as well as the levels of security or vulnerability which they contain. The results could then further be incorporated into a simulation of the relevant processes and multiple scenarios can be run in order to determine the likelihood of the potential challenges coming to pass, and the likely outcomes of the identified challenging situations should they come to pass.

Further operational analysis and simulation exploration may then lead to suggested improvements or additional procedures which may then, under particular circumstances, be used as a tool to support decision making regarding current and future security screening practices and procedures.

6. Value-Added

CQADS consultants bring additional benefits to the projects in which they are involved; in the context of providing services to CATSA, these benefits are mostly due to the proposed team members’

* affiliations to relevant professional and academic organizations,
* teaching activities in relevant Subject Areas,
* ability to obtain and hold security clearances, and
* expertise in more than one Subject Area,

Membership and involvement in professional and academic organizations provides an external (and independent) corroboration of a resource’s expertise and abilities:

* **Yiqiang Zhao** is a member of the Institute for Operations Research and Management Sciences (INFORMS), a Past President of the Probability Section of the Statistical Society of Canada (SSC), the director of the Laboratory for Research in Statistics and Probability (LRSP) since 2004, a co-program chair of the Canadian Operation Research Society (CORS), a member of the Canadian Applied and Industrial Mathematical Society (CAIMS) and a member of the Canadian Mathematical Society (CMS).
* **Kevin Cheung** is a member of CORS, serving as the Optimization Cluster Chair for their 2014 Annual Conference.
* **Jason Nielsen** is a member of the SSC and he has successfully completed exams of the Society of Actuaries (Calculus, Probabilities).
* **Mohamedou Ould Haye** is a member of the SSC, and he has organized International Conferences in Probabilities and Statistics, with support from the Fields Institute.

Teaching advanced quantitative courses requires a deep and thorough understanding of the theory underlying the Subject Areas (and some of their cousins), demonstrating yet another layer of competence: the proposed team members have been teaching advanced courses in Probability Theory, Statistics, Statistical Modeling, Statistical Computing, Time Series Analysis, Stochastic Processes and Queueing Theory, Data Mining, Linear and Combinatorial Optimization, Analysis, and Differential Equations at various Universities:

* **Patrick Boily** has taught over35 courses at the University of Ottawa and l’Université du Québec en Outaouais.
* **Yiqiang Zhao** has taught over 40 courses at Carleton University, Queens University and the University of Winnipeg.
* **Kevin Cheung** has.taught over 20 courses at Carleton University. He is also the recipient of the Carleton University’s Excellence in University Teaching with Technology Award for 2013 and one of ten finalists of the YouTube Next EDU Guru Contest and winner of the 2012 Khan Academy Prize for hist YouTube Channel MathApptician.
* **Jason Nielsen** has taught over 15 courses at Carleton University.
* **Mohamedou Ould Haye** has taught over 10 courses at Carleton University.

Additionally, all these team members have been invited to give lectures or conferences talks on their research at prestigious international conferences.

As an institution, Carleton’s holds Reliable Status; to CQADS’ knowledge, individuals are thus prohibited to hold a security clearance at a higher level through the University. However, Patrick Boily and Jennifer Schellinck both have held Secret Clearance (Level II) status in the past, the clearance residing with previous employers. Jennifer is currently in the process of migrating the Level II Security Clearance back to Sysabee (thus ultimately resting in her hands), while Patrick’s Level II Clearance can be duplicated by prospective clients (this was recently successfully completed by Apption, an Ottawa analytics company). Security clearances for other team members have been initiated.

Finally, it should be noted that some of our team members meet the qualification requirements of more than one Subject Area, illustrating the notion that the domains are inter-connected, and that it is difficult to be an SME in one without also being an SME in some of the others. This provides an additional layer of security for CATSA should one of the SME find themselves indisposed for any period of time.