# Inspecting less to inspect better: The use of data mining for risk management by customs administrations

Anne-Marie Geourjon // Bertrand Laporte //
Ousmane Coundoul // Massene Gadiaga

⮕ Anne-Marie Geourjon, Ferdi.

⮕ Bertrand Laporte, Cerdi, Université d'Auvergne.

⮕ Ousmane Coundoul, Douanes Sénégalaises.

⮕ Massene Gadiaga, Douanes Sénégalaises.

**Abstract**

In order to limit the number of intrusive inspections, the most modern customs administrations rely on risk analysis as the only effective tool for facilitating trade and securing their operations given the important growth in trade volume in recent years. However, customs administrations in developing countries have been slow to adopt this methodology. This paper demonstrates that based on the experience of Senegal, data mining and statistical scoring techniques can be used effectively by customs in developing countries to assess risk and to assign declarations to the various inspection channels. The paper also shows that the in-house development of this type of system by customs administrations in these countries significantly advances the process of modernization.

## 1. Introduction

Under the revised Kyoto Convention on the simplification and harmonization of customs procedures, the WCO recommends that intrusive customs inspections be limited.[1] This proposal has also been discussed in the context of WTO trade facilitation negotiations.[2] Therefore, despite the temptation to ramp up systematic inspections in the light of the events of September 11, 2001, the most modern customs administrations have continued to rely on risk analysis as the only effective means for both facilitating trade and securing their own operations, given the important growth in trade volume in recent years (Harrisson, 2007).

It is also necessary for the customs administrations of developing countries to have a structured approach to risk analysis when determining how to process a particular trade transaction (Walsh, 2003; Widdowson, 2005). These countries are simultaneously confronted with the growth in trade flows and the demands of private operators as well as by the pressures placed on them by governments to secure revenue mobilization. However, they have been slow to move in this direction and to implement the latest risk analysis and risk management techniques (Geourjon & Laporte, 2005; Geourjon, Laporte, & Rota Graziozi, 2010; Hintsa *et al.*, 2011).

The information systems used by customs administrations in developing countries screen declarations by defining and applying iterative selection rules based on criteria that are largely qualitative and dual, in addition to which random targeting is also applied. The traditional selection methods used in the electronic customs clearance systems of these countries remain very much dependent on human judgment, which represents a major shortcoming, given moral hazard. These systems are also static and rigid as the rules defined in them are seldom updated, allowing fraudsters to modify their behavior to avoid detection.

Private inspection companies working in developing countries offer customs administrations risk analysis services as part of their contracts with governments to carry out pre-shipment and destination inspection and/or scanning services. The systems offered by these companies are standardized and based on their own data. Customs administrations therefore have difficulty adapting them to their purposes and in practice seldom use them to inform the selection process for their inspections, which essentially continues to be based on their own traditional methods. The main reason for this situation is that the risk analysis services offered by the private inspection companies and the import verification programs they were contracted to provide have two conflicting objectives. Regarding risk analysis services, the aim is to modernize administration procedures, which implies a partnership between customs and the private company, whereas in import verification, the effectiveness of the double checking system relies on customs and the private company being independent of one another in order to avoid any hint of collusion. It is therefore incompatible to include both inspection services and risk analysis services in the same contract. Moreover, the companies that offer

---

1 WCO, 2003.
2 Article VIII of the GATT, which aims to limit the number of procedures required for import and export operations.

the latter base this service purely on the data collected during the course of their own inspections… (Dequiedt, Geourjon, & Rota-Graziosi, 2009, 2012).

In the majority of developing countries and particularly in Sub-Saharan Africa, customs administrations continue to carry out intrusive inspections on large numbers of containers, resulting in the proportion of detected incidences of fraud generally being less than 3% (as is the case in Benin, Côte d'Ivoire, Mali, and Senegal, for instance), hence the need for customs administrations to develop their own risk analysis and management systems that will be modern, effective, and based on customs data in order to be able to inspect less but more effectively. The idea is to apply the risk analysis methods used in many other sectors (banking, insurance, security, etc.) and to adapt them to the context of customs. In practice, if risk analysis is to be applied in every sector and in every organization, this will require a specific approach in each case (Gates, 2006).

The aim of this paper is to show that data mining and relatively simple statistical scoring methods can allow customs administrations in developing countries to assess risk and hence to effectively limit inspections. This paper also shows that the development of these techniques will contribute to the modernization of these administrations. The following section describes some of these techniques; Section 3 outlines how they have been implemented in West Africa and more specifically assesses their use in Senegal; Section 4 analyses the impact of the use of these techniques on the modernization process; and the final section summarizes the main conclusions.

## 2. Statistical scoring techniques for measuring risk: An innovative tool for customs

For customs administrations, the goal is to design a system that will assist them in effectively deciding which transactions to inspect. Although already in use in many other sectors, statistical scoring techniques are seldom used by customs administrations in developing countries. However, the advantages of these techniques are far greater than simply carrying out valuation and/or tariff classification controls. A recent study by Grigoriou (2011) shows the advantages of statistical scoring for carrying out inspections with a view to enforcing technical, sanitary, and phytosanitary regulations.

### *Customs information systems: Information flows that need to be organized*

All risk analysis depends on information, which needs to be available and processed correctly. The main obstacle to developing systems based on data mining is the absence of reliable data on detected incidences of fraud. This may be due to one of two reasons: weaknesses among customs administrations in litigating offenses, and the poor traceability of data relating to this litigation (customs violation reporting). Nonetheless, significant information flows concerning customs fraud are available and should be processed in order to analyze overall risk and to manage it. Figure 1 below distinguishes between data on confirmed and suspected fraud. The use of risk analysis and management in order to assign declarations to the various inspection channels depends mainly on the use of existing historical data on detected incidents of fraud over a given period. The results obtained are completed by using available information relating to suspected fraud (inspection criteria) and

by the use of a given percentage of random inspections. The effective operation of the system depends on the quality and use of all of these data and therefore leads to the setting up of a data platform that gathers information on detected customs violations and suspected fraud linked item by item to the components of the declarations in question. This platform forms the heart of the overall risk analysis and management system and will facilitate targeting for second-line inspections and assessment of customs violation.

*General architecture of the system: Four complementary approaches*

The system for targeting declarations for customs inspection should take into account the most important components of a trade transaction, which are: the origin of the goods and the trade channels through which they have passed; the goods that are the object of the trade transaction; and the operators involved in the transaction. The origin and trade channels are important because they may reveal abnormal channels, as determined by knowledge of the most usual and regular trade transactions. The type of goods, which determines taxation rates, restrictions, prohibitions, restrictive measures, qualitative controls, and special tariff arrangements together with the customs value are two key factors in the presumption of fraud. Finally, the operators involved in the transaction form the remaining component of the system. Although the importer is the main figure involved, there are other operators in the chain (exporter, shipping company, banker, forwarder, and so on).

Based on these components, it is possible to design a targeting method by combining four different approaches based on statistical analysis of detected cases of fraud and on the assessment of incidents of suspected fraud by customs officials, particularly based on intelligence gathering activities. These four approaches are applied iteratively in order to assign declarations to the appropriate inspection channel.

**Figure 1: Information flows on customs fraud theoretically available to customs administrations**
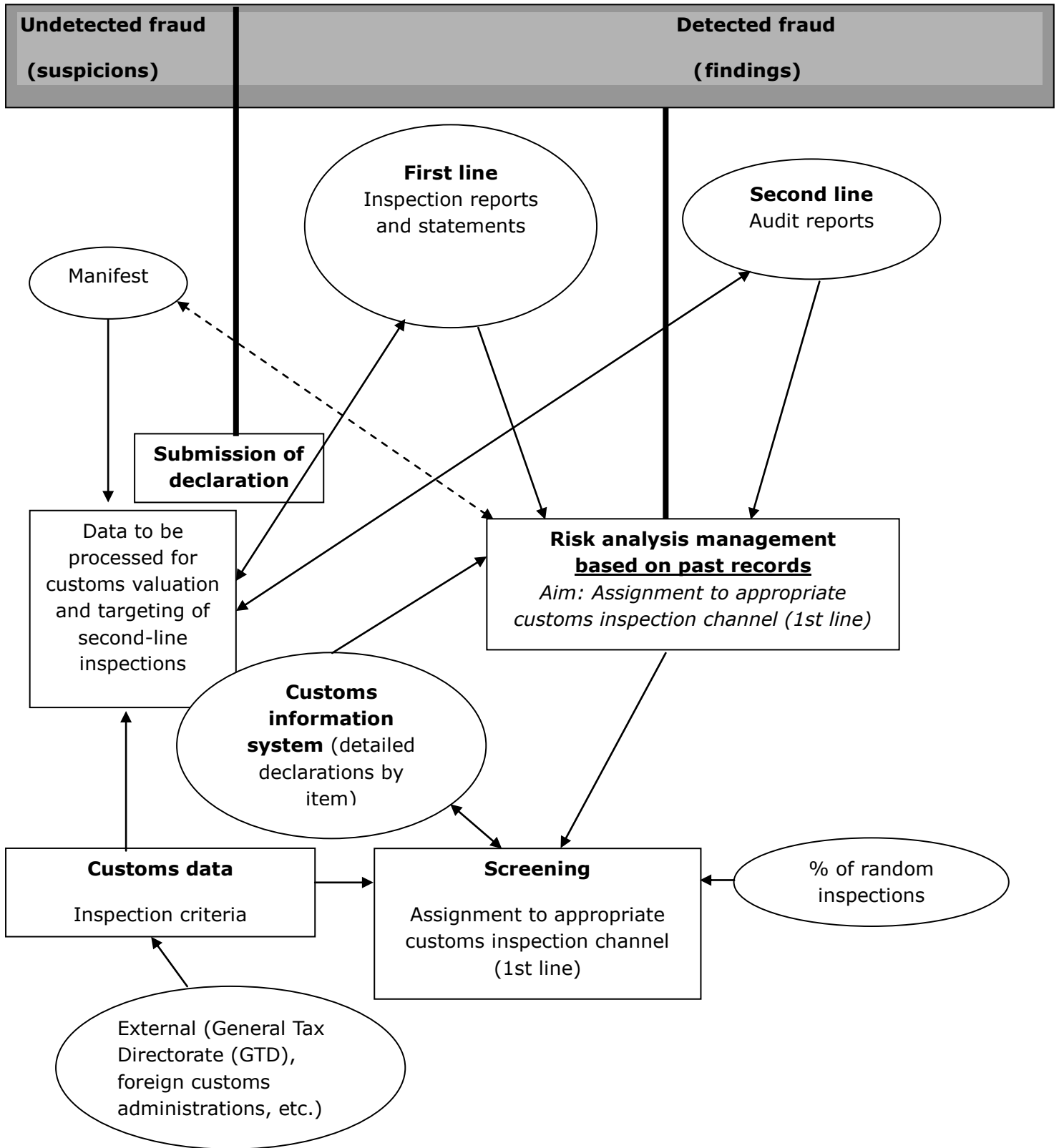
*The first approach involves verifying every new operation* that concerns an operator, a type of goods, or a trade channel for which the customs database does not supply any information and that it would therefore be advisable to single out based on past knowledge. Crucially, this systematic inspection should prompt the operators to identify themselves correctly, which is essential to customs administration procedures (as well as to tax administration) since post clearance controls should gradually replace import controls at the front line.

*The second approach, which is just as radical as the first, involves the systematic inspection of transactions based on factors linked to specific characteristics* (for example, transactions to a value greater than x currency units, the fact that the operator has not undergone an inspection in the past x weeks, and so on) *and based on suspicion of fraud (unconfirmed fraud).*

*The third approach results from an examination of the documents making up the import dossier and the statistical analysis of confirmed cases of fraud in order to deduce information on the risk of fraud*: each criterion identified (relating to trade channels, goods, and operators) is assigned an individual score based on the available statistical data. A representative overall score indicating the fraud risk of the transaction to be inspected is then obtained by combining the individual scores of the various criteria applied.

*The fourth approach involves a purely random selection.* It allows customs to limit the number of inspections but particularly to prevent economic operators and/or unauthorized state officials from modifying their behavior on the basis of statistical information defining the criteria used to detect fraud in the third approach. A significant proportion of the transactions that are singled out for inspection should therefore be selected randomly, especially given that the statistical analysis will initially be based on customs violation data that will be fairly unreliable due to moral hazard and information asymmetry.

The third approach forms the heart of the method because it allows customs to determine the probability of fraud for each transaction based on objective risk criteria identified using statistical inferences. A first step will be to determine the fraud criteria based on an ex-post analysis of confirmed cases of customs violations. A second step will be to apply these criteria to each new transaction in order to determine the probability of fraud (overall score) and, ultimately, the inspection level. Prerequisite to any ex-post analysis is the existence of a database of customs violations.

It is important to emphasize that the system's performance depends on the appropriate use of the second and third approaches. The second approach is derived from the analysis of information on fraud that has not yet been confirmed but that has been perceived or suspected by the customs officials in charge of control selectivity. It should allow for coverage of new risks of fraud. The third approach is built on a historical analysis by statisticians of confirmed cases of fraud and assessing all known risks over a reference period. Assigning too much importance to the second approach and multiplying the number of criteria that lead to systematic inspection cancels out the benefits derived from carrying out scientific risk analysis.

*Statistical analysis of confirmed cases of fraud (Laporte, 2011)*

In order to accurately target the declarations that carry a risk of fraud, data analysis first needs to be carried out. This analysis will involve identifying the characteristics of declarations from the recent past that have involved fraud (a customs violation) and to detect statistical regularities in these incidents of fraudulent behavior. All available information will be used: the content of the certificate of verification (CV) produced by the companies in charge of the import verification program, the manifest, the detailed declarations, and the inspection reports (first and second line inspections) over the reference period. The statistical regularities identified will allow risk profiles to be outlined.

Although the data are largely qualitative when using traditional methods of selectivity, statistical analysis allows the establishment of a quantitative risk scale. For example, in order to measure the profile of importers, the frequency of violations will be calculated for each importer (the ratio of the number of fraudulent declarations by a given importer to the overall number of declarations submitted by that importer over a given period). Importers are then placed on a scale of 0 to 1 (or 0 to 100): 0 for importers that represent no fraud risk and 1 for importers associated with a high risk of fraud. This type of calculation can be made for all potential risk criteria including trade channels, operators, and goods, and it can allow the determination of risk profiles for each criterion.

*Assignment of a declaration to a customs clearance channel*

The risk profiles obtained are combined to inform decision making with regard to which customs clearance channel a particular declaration should be assigned. The aim is to give each new declaration a score obtained by combining the fraud rates for each of the various criteria. This score should best reflect the risk of fraud, or rather the probability of fraud being committed. Assignment to one of the customs clearance channels is based on this score and on thresholds previously determined using statistical analysis.

With the simplest systems, the declaration's score can be obtained by taking a simple or weighted statistical average of the fraud rates (risk profiles) of the various criteria used or simply by taking the highest value from among the criteria (although other combinations can be used). Prior to this, the most significant criteria will have been determined in an ad hoc fashion by customs officials responsible for the inspection activities or by statistical trial and error to arrive at the best combinations. The most commonly used criteria are the importer, the shipping agent, the Harmonized System (HS) position, the customs regime, the country of provenance, and the country of origin of the goods. Criteria may also be combined.

More elaborate systems use statistical distribution properties to effectively combine customs data. Econometric models also allow the identification of the risk criteria that best account for an act of fraud and to calculate the probability of fraud for each new declaration. This probability is then the resulting score for that declaration. To do this, the following equation first needs to be solved based on the history of declarations:

$$\Pr\left[Fraud_{ij} = 1\right] = \alpha + \beta_1\, fq_{criterion1_{ij}} + \beta_2\, fq_{criterion2} \ldots_{ij} + \beta_N\, fq_{criterionN_{ij}} + \varepsilon_{ij}$$

Where Pr is the probability; $Fraud_{ij}$ the binary variable 0/1 for the declaration i, product j (1 if fraud and 0 if no fraud for declaration i for product j); $fq\_{ij}$ the frequency of fraud for each risk criterion associated with declaration i and product j; $\varepsilon$ the random deviation (that which cannot be explained by the criteria included in the equation); and $\alpha$ and $\beta$ the parameters of the equation to be solved.

**3. Experience in West Africa: The case of Senegal**

With the assistance of the IMF's West AFRITAC, five countries in West Africa are currently developing this type of system: Benin, Burkina Faso, Côte d'Ivoire, Mali, and Senegal. The system has only recently been introduced in Benin and Mali and is in a testing phase in Côte d'Ivoire and Senegal. In Burkina Faso, it is only just being launched.

Each country has adapted the system to its own context (types of operators, integrated customs clearance systems, involvement of anti-fraud services, etc.). Thus Benin was able to develop a statistical scoring system within ASYCUDA++ based on an econometric analysis of fraud criteria and their various combinations thanks to technical support funded by external sources. Mali has circumvented the difficulties associated with a closed computer system (ASYCUDA++) by using statistical risk profiles to assign a risk category (low, medium, or high) to operators, HS positions, and country of origin in its transitional risk analysis and management system. It then combines these risk categories with simple rules to direct declarations to a particular customs clearance channel (for instance, a declaration is assigned to the red channel when two criteria are "high risk"). Côte d'Ivoire has developed a transitional system that currently functions in parallel with the SYDAM-World screening system. When a declaration is assigned the maximum score for the three chosen criteria (importer, HS position, and provenance), it is directed to an inspection channel. The risk analysis is carried out based on the DPI (Declaration Prior to Import), hence prior to submission of the detailed declaration, which allows inspection services to anticipate the need for inspections. During the course of the first quarter of 2012, Senegal aims to integrate into GAINDÉ (its existing system for automatic management of customs information and trade) a dual/two-track system called SAGAR that uses two broad categories of importers: registered and unregistered. Risk analysis for the registered operators takes place based on the fraud rate for four criteria: importer, country of origin, customs regime, and HS position. Unregistered (occasional) operators are systematically sent for detailed inspection.

*Customs risk management previously applied in Senegal*

Customs authorities in Senegal have been applying a system called SIAR (*Système Informatisé d'Analyse de Risque* – which stands for "Computerized Risk Analysis System), which was developed by the private company COTECNA to select declarations for inspection. A steering committee for SIAR, consisting of representatives from COTECNA and the Senegalese customs administration, meets regularly to adapt the system to needs on the ground.

In practice, two risk management systems coexist. The first, SIAR-Senegal, is based on the analysis of data from the PVI (*Programme de Vérification des Importations* – Import

Verification Program) and determines which import shipments should undergo pre-shipment inspection. Based on a certificate of verification, the shipments are assigned to one of five inspection channels for destination inspection. The second relates to non-PVI imports, which are processed by the Senegalese customs administration's GAINDÉ system, which works on the basis of simple inspection criteria. Thus 70% of imports bypass SIAR-Senegal and are therefore not subjected to the COTECNA risk analysis.

*SIAR COTECNA-Senegal*

Imports with an FOB value of less than CFAF 1 million are excluded from the PVI, which means that importers do not have to submit a DPI. A DPI must be submitted for imports with a CIF value of under CFAF 3 million (and over CFAF 1 million FOB), but they need not undergo pre-shipment inspection. For DPIs with a value of over CFAF 3 million, an inspection is carried out before loading. The certificate of verification is then electronically submitted to Senegalese customs using the GAINDÉ system.

SIAR is organized into two levels: Upstream SIAR (*SIAR-amont*) and Downstream SIAR (*SIAR-aval*) Upstream SIAR determines the type of intervention undertaken by COTECNA before the goods are loaded. The process depends on the analysis of the DPI and assigns imports to one of three channels:
- The blue channel: goods exempted from intervention by COTECNA;
- The green channel: providing a documentary check for the tariff classification and the price analysis;
- The red channel: implying a physical inspection of the goods before loading.

The aim of Upstream SIAR is to limit the number of physical pre-shipment inspections to 10% of import transactions.

Downstream SIAR determines the type of intervention undertaken by Senegalese customs on arrival of goods that have a certificate of verification. It uses five channels:
- The blue channel: automatic issue of release warrant without any inspection, particularly for goods that have undergone a physical pre-shipment inspection;
- The green channel: documentary check;
- The yellow channel: inspection on importer premises;
- The orange channel: documentary check along with inspection by scanning;
- The red channel: physical inspection of goods.

Assignment of the declarations to one of the channels whether by Upstream SIAR or Downstream SIAR is done based on a score (econometric model) given to the transaction as well as on the inspection criteria (specific tariff positions, operators, etc.). COTECNA uses its own database, which consists of the results of its own inspections.

*How risk is taken into account in GAINDÉ*

For transactions that fall outside the PVI and are therefore not subjected to the SIAR risk analysis, the GAINDÉ system directs the declarations to one of the five inspection channels purely based on the criteria defined by the SIAR steering committee on the basis of its perception of fraud risk:
- The blue channel: changes and goods taken from warehouses;

- The green channel: declarations for perishable products, some suspensive regimes, and importers to whom special arrangements apply;
- The yellow channel: heavy or dangerous goods as well as personal belongings;
- The orange channel: FLCs (fully-loaded containers) that fall outside the PVI;
- The red channel: exemption regimes, products susceptible to fraud, and declarants, importers, and blacklisted countries of origin, or when the value of the goods is higher than a fixed amount as defined by each customs office.

There are general characteristics common to all the customs offices as well as local criteria used by individual customs offices, especially for the red channel.

The outcome of all of the controls (SIAR and customs) should in principle feed into SIAR in order to update in real time the risk profiles that form the basis of the statistical scoring. However, this information exchange does not take place, which greatly reduces the effectiveness of this type of system.

### *Towards a modern risk analysis and management system integrated into GAINDÉ: SAGAR*[3]

Since 2009, the Senegalese customs administration has been working to develop its own system of risk analysis and management independent from COTECNA's SIAR and based on its own data.

*Prerequisites*

The development of a customs risk analysis and management system implies the need for a number of prerequisites, which Senegal customs has gradually put in place. These relate to their computer system, the institutional framework, and the availability of the data and resources that need to be released. Unlike countries that use ASYCUDA 2.3 or ++, GAINDÉ is a system developed by Senegalese customs and is therefore open, which facilitates the in-house development of a risk analysis and management application. The identification of the risk analysis project as being of high priority to Senegalese customs' modernization plans led to the creation of a favorable institutional framework, namely an analysis and decision-making support bureau (*Bureau d'Analyse et d'Aide à la Décision* – BAD), consisting of a customs inspector (Bureau Chief), two statisticians/econometricians, and IT professionals along with the creation of a risk management committee consisting of representatives of the BAD and the Intelligence and Documentation Bureau (*Bureau du Renseignement et de la Documentation* – BRD). Once constituted, the BAD was able to work on the compilation of a coherent database before conducting the statistical study necessary for the development of the system. Finally, the means of execution were put in place for the project both in terms of human resources and financing.

---

3        Automatic Risk Management and Analysis System (*Système Automatisé de Gestion et d'Analyse de Risque*).

*SAGAR: Initial version*

SAGAR combines the four complementary approaches described above. Representing around 40% of import transactions, the importance of unregistered operators in Senegal led the Senegal customs administration to design a two-track risk analysis and management system (Figure 2) whereby importers are divided into two groups, with a different risk management strategy applied to each group. However, the statistical analysis cannot be applied to unregistered operators because the operator criterion, which plays a major role in determining the overall risk score, does not allow these operators to be distinguished from one another.
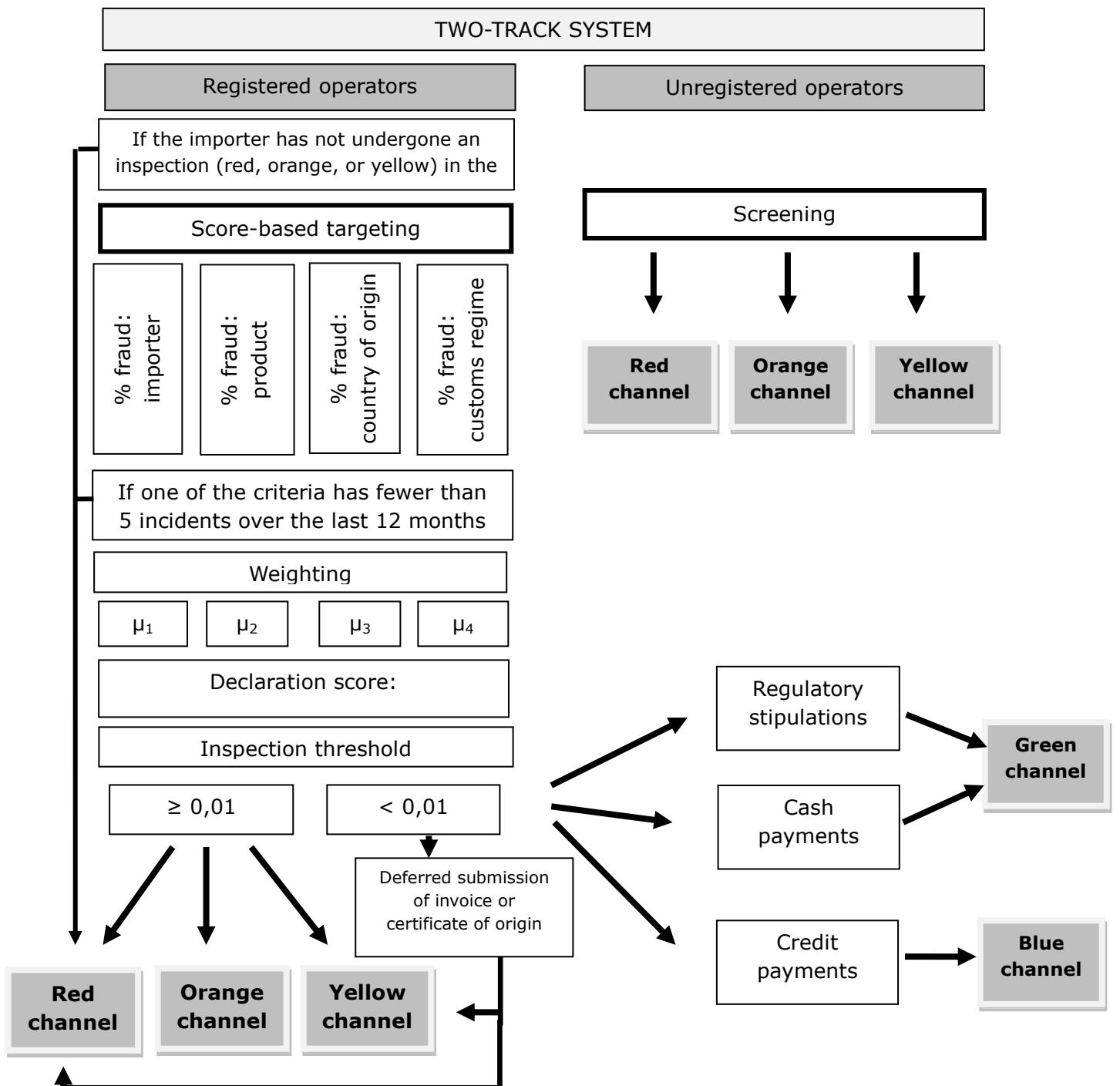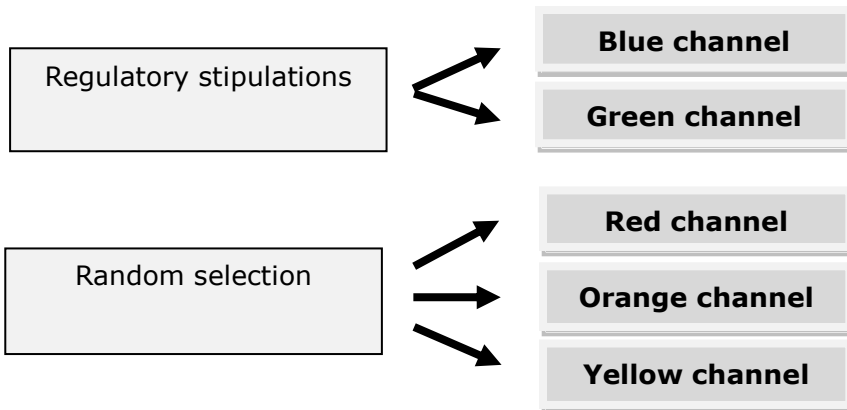
The five existing clearance channels are retained in the new system. Version 1.0 of SAGAR is based on a simple combination of the risk profiles of the four criteria: the importer, the country of provenance, the product (HS code), and the customs regime. Version 2.0 will use econometric analysis to establish which criteria are used and in which combination.

*Evaluation of the initial version of SAGAR: Encouraging first tests*

It is difficult to evaluate the performance of a new risk analysis and management system because a reference situation to which the targeting outcomes can be compared first needs to be defined. Two reference situations are conceivable: the first would be to compare the targeting outcomes of the new system to those of the existing system, while the second would be to compare the outcomes of the new system to a random selection. However, in either case, it is not easy to carry out the necessary tests because this requires that both systems function in parallel, something that is not possible with the integrated customs clearance systems used in most countries. The option that is often chosen is therefore the first: comparing the original system with the new system. After all, the aim of the tests is to establish whether, despite insufficient customs data, the new system is at least as effective at detecting fraud as the old, while inspecting less. What is needed therefore is to compare the outcomes of the screening and inspections of the two systems not in real time but over a given period in the past, with the performance of the new system determined a posteriori. The question that the tests answer is therefore the following: Were the violations detected by the old system also detected by the new system but with a lower inspection rate?

The tests carried out for SAGAR are based on a calibration of the system over four consecutive quarters. This involved determining the risk profiles for each criterion and the rules for combining them over a given period. The system thus established is then applied to the next quarter. In the case of Senegal, the system was calibrated based on the year 2010 and applied to the first quarter of 2011, then calibrated based on the last three quarters of 2010 and on the first quarter of 2011 and applied to the second quarter of 2011. As the full programming for SAGAR has not yet been finalized, SAGAR only uses the statistical approach to determine the risk category of declarations, and only for registered operators.

**Figure 2: Decision tree envisaged by SAGAR for the categorization of declarations**

```
Regulatory stipulations  →  Blue channel
                         →  Green channel

Random selection  →  Red channel
                  →  Orange channel
                  →  Yellow channel
```

**TWO-TRACK SYSTEM**

| Registered operators | Unregistered operators |

**Registered operators:**

If the importer has not undergone an inspection (red, orange, or yellow) in the

**Score-based targeting**

- % fraud: importer
- % fraud: product
- % fraud: country of origin
- % fraud: customs regime

If one of the criteria has fewer than 5 incidents over the last 12 months

**Weighting**

$\mu_1$  $\mu_2$  $\mu_3$  $\mu_4$

Declaration score:

Inspection threshold

≥ 0,01    < 0,01

→ Red channel
→ Orange channel
→ Yellow channel

Deferred submission of invoice or certificate of origin

**Unregistered operators:**

Screening
→ Red channel
→ Orange channel
→ Yellow channel

Regulatory stipulations → Green channel
Cash payments → Green channel
Credit payments → Blue channel

As the tariff classification is one of the risk assessment criteria, the data are by necessity organized by tariff headings in the database. Hence, if a customs declaration consists of three items, the declaration is divided into three distinct transactions and an overall score is assigned to each of these transactions. The risk classification having been carried out based on the declaration, the classification results are then presented at the declaration level, with the overall risk score assigned to the declaration being equal to the highest score from among individual transactions.

During the course of the first quarter of 2011, all declarations[4] (7,947 in all) were inspected (red, orange, yellow, or green channels) and directed to a channel either based on Downstream SIAR or by using GAINDÉ's criteria. Only 56 of the inspected declarations (or 0.7%) were subject to litigation. During the course of the second quarter, 7,633 declarations were inspected (red, orange, yellow, and green channels), or 99.8% of declarations. A total of 60 declarations (or 0.8%) were subject to litigation (see Table 1). This very low rate of litigation for a high inspection rate is justification enough for the development of SAGAR with the aim of improving the effectiveness of risk analysis and management by Senegal customs.

**Table 1: Frequency of detected fraud among registered operators**

|  | 1st quarter | | 2nd quarter | |
|---|---|---|---|---|
| Litigation | Number of declarations | % | Number of declarations | % |
| No | 7,891 | 99.3 | 7,573 | 99.2 |
| Yes | 56 | 0.7 | 60 | 0.8 |
| Total | 7,947 | 100 | 7,633 | 100 |

As the application of SAGAR in GAINDÉ has not yet been finalized, only statistical screening has been taken into account in the tests (Approach 3 of the system). It only distinguishes two channels: strict inspection channels (red, orange, and yellow) and other channels. The results outlined in Table 2 therefore only reflect the channels common to both systems, i.e., the detailed inspection channels for which the outcomes are known.

**Table 2: Assignment to strict inspection channels (red, yellow, orange) and outcomes of inspections**

|  | 1st quarter | | | 2nd quarter | | |
|---|---|---|---|---|---|---|
|  | Litigation | | | Litigation | | |
|  | No | Yes | % | No | Yes | % |
| SIAR+GAINDÉ | 5,192 | 34 | 0.65 | 4,812 | 45 | 0.93 |
| SAGAR | 2,004 | 28 | 1.38 | 1,368 | 24 | 1.72 |

---

4 No declarations were assigned to the blue channel during this period.

In the first quarter of 2011, Downstream SIAR and GAINDÉ assigned 5,192 declarations to the red, orange, or yellow channels. Among them, only 34 were subject to litigation, a litigation rate of 0.65%. SAGAR would only have assigned 2,004 declarations to the same channels, i.e., a 60% reduction in inspections. Among those, 28 of the 34 declarations that were subject to litigation were targeted, which would have produced a litigation rate of 1.38%. In the second quarter, SIAR and GAINDÉ assigned 4,812 declarations to the red, orange, and yellow channels, of which 45 were subject to litigation, or a litigation rate of 0.93%. SAGAR would only have assigned 1,36 declarations to the same channels, i.e., a reduction in inspections of more than 70%. Among those, 24 of the 45 declarations that were subject to litigation were targeted, producing only a slightly worse result than in the first quarter, during which the litigation rate would have been 1.72%.

These results are particularly interesting. In the first quarter, SAGAR detected more than 80% of the declarations that were subject to litigation while reducing the number of inspections by a factor of more than 2.6. Moreover, this was achieved with SAGAR's targeting mechanism, which uses only one of the four approaches making up the risk analysis and management system. Taking into account the other three should therefore allow it to identify virtually all the cases of fraud revealed by SIAR and GAINDÉ but with a significantly reduced number of inspections. Moreover, reducing the number of inspections should allow for an improvement in the quality of the inspections carried out and hence the detection of additional cases of fraud.

**4. Data mining: An accelerator for customs modernization**

The experiments carried out in West Africa have helped to address some doubts with regard to the usefulness of this type of technique in administrations with limited resources but which are already undertaking far-reaching reforms. This raises two questions. Firstly, can this type of sophisticated system be developed in these administrations? Secondly, is it appropriate to dedicate time and resources to the development of data mining and scoring techniques far removed from the purely customs-related concerns of the numerous activities already undertaken (combating fraud, assessing value, etc.)?

First of all, the weakness of customs administrations does not represent an obstacle to the development of this type of system. On the contrary, some aspects of the dysfunctionality within these administrations, including the lack of ethics, are directly addressed by the use of scientific risk analysis techniques that remove the need to rely on decision-making based on human judgment. In addition, most customs administrations already have the skills necessary to apply these techniques, although capacities related to data mining and statistical inference may need to be strengthened.

Investing in the development of these techniques significantly contributes directly or indirectly to the modernization process in the customs administrations that choose to adopt them.

### *Direct effects on the reform process*

It itself, risk analysis provides powerful leverage for the general reform of customs administrations particularly because it requires closer cooperation between the various services in charge of inspections and intelligence. Limiting the number of first line inspections also frees up inspectors, who can be redeployed to carry out post clearance audits. The development of this type of inspection, which has been virtually non-existent until now, represents an important component of any reform program.

The prerequisite for using data mining to assess the risk of fraud is the compilation of a reliable database of customs violations. Since it is essential to ensure the traceability of instances of detected fraud at every level (both first and second line), this implies computerizing customs litigation, i.e., making it possible for records to be created online and fines to be issued electronically. Customs administrations that have adopted this approach have been led to completely review all their litigation procedures before computerizing them, which allowed them to simplify procedures, and this drew attention to a number of profound dysfunctions. The process also represented an opportunity to bring together IT specialists, statisticians, and customs inspectors and to have them work together through a rigorous approach conducive to the adoption of new technologies.

### *Structural change in the culture of customs administrations*

The use of data mining brings about a significant change in attitudes and behavior. One of the foundations of a modern customs administration is information. If customs administrations gather or have access to a large amount of data, these are often scattered, compartmentalized, incomplete and, as a result, difficult to mine. Moreover, when data exist, they are generally poorly utilized or not used at all. The approach taken by the customs authorities that took part in these experiments created an awareness of the importance of using data over and above the specific aim of developing a risk analysis and management system. It allowed these administrations to realize the value of specific tasks, including the production of monitoring indicators. This change in culture can only facilitate the achievement of the modernization program's objectives, particularly with regard to management and human resources management.

### *A doorway to other projects*

The development of risk analysis and management systems using data mining should open up new possibilities for other innovative modernization projects, including the categorization of operators and the introduction of performance contracts, for instance.

Defining the main categories of the customs administration's clients is essential to adapting procedures and controls especially when the aim is to make these more efficient. The analytical results obtained with regard to the importer risk form a natural part of the criteria that need to be met in order for these importers to be granted the status of "authorized economic operator," as advocated by the WCO and the WTO. The results are also useful in

establishing lists of high-risk operators when combined with other criteria.[5] The categorization of operators then allows for the adaptation of screening procedures based on risk. In theory, authorized economic operators are never directed to the red channel except through random selection, in which case they are obliged to submit their goods to inspection, though this will happen only very rarely. Operators considered very high risk are almost always directed to the red channel because being classified as high risk represents an inspection criterion. It is for the other operators that risk analysis and management plays the most decisive role when a decision is made as to whether or not to direct them to the red channel (around 20% of them are so directed). Documentary checks are traditionally carried out for products that are not assigned to the red channel and for which regulations require that specific documents be produced.

Recently, the customs administration in Cameroon launched a pilot program for performance contracts that had very positive outcomes with regard to reducing customs clearance delays, mobilizing revenue, and improving officials' behavior (Cantens, Raballand, Strychacz, Tchouawou, 2011). These performance contracts aim to promote a results-based culture in the administration, outlining objectives for which the expected results are established. This new approach depends on the quantification and evaluation of indicators, which requires access to the necessary data. The development of risk analysis systems based on data mining, which implies the availability of a set of data on fraud, is based on the same philosophy and should open the way to this type of experiment in the customs administrations of other African countries.

## 5. Conclusion

Risk analysis is indispensable for customs administrations in developing countries if fewer inspections are to be carried out while inspections become more effective. These administrations have recently come to realize the possibilities offered by data mining thanks to initiatives by the inspection companies that proposed using the results of the tools they developed for their own screening procedures. However, customs administrations have not been able to appropriate these tools and they now feel the need to develop this type of system in-house. Over the past five years, five countries in West Africa have launched projects in this regard with the support of West AFRITAC and the IMF. Each administration has adopted an approach that was tailored to its context and its needs, and significant progress has been achieved. Based on the case of Senegal, a comparison between the targeting results of the system using data mining to assess risk and those of the traditional screening system demonstrated the progress made with regard to trade facilitation. In addition, these experiments show the positive impact of these projects on the reform process by means of their direct effect on work procedures and methods, the development of an information culture, and the opening of the way toward other innovative projects.

---

5 In addition to occasional operators, high risk operators can be identified by analyzing the file of taxpayers in order to identify suspicious operators (operators that match several names or from whom tax has not been collected on any activity for the past fiscal year, for instance).

## Bibliography

Cantens T., G. Raballand, N. Strychacz, & T. Tchouawou (2011), "Réforme des douanes africaines: Les résultats des contrats de performance au Cameroun," *Afrique-Notes de Politique commerciale*, Note #13, Washington, DC, World Bank.

Dequiedt V., A-M. Geourjon, & G. Rota-Graziosi (2009), "Les programmes de vérification des importations à la lumière de la théorie de l'agence," *Afrique Contemporaine*, 230, pp. 151-166.

Dequiedt V., A-M. Geourjon, & G. Rota-Graziosi (forthcoming 2012), "Mutual supervision in preshipment inspection programmes," *Journal of Development Economics*.

Gates S. (2006), "Incorporating strategic risk into enterprise risk management: A survey of current corporate practices," *Journal of Applied Corporate Finance*, Vol. 18, No. 4, pp. 81-90.

Geourjon A-M. & B. Laporte (2005), "Risk management for targeting customs controls in developing countries: A risky venture for revenue performance?" *Public Administration and Development*, Vol. 25, pp. 105-113.

Geourjon A-M., B. Laporte, & G. Rota-Graziosi (2010), "Comment moderniser l'analyse du risque et la sélectivité des contrôles douaniers dans les pays en développement?" *OMD Actualités*, No. 62, pp. 29-31.

Grigoriou C. (forthcoming 2012), "How can risk management help enforce technical measures?" in O. Cadot & M. Malouche (eds), *Non Tariff Measures: A Fresh Look at Trade Policy's New Frontier*, Washington, DC/London, World Bank/CEPR.

Harrisson M. (2007), "Challenges for Customs and Supply Chain Security: 'The Demise of Risk Management?'," Annual Conference on APEC Centers, April 18-20, Melbourne, Australia.

Hintsa J. *et al* (2011), *Customs Risk Management: A Survey of 24 WCO Member Administrations*, Cross-border Research Association, EPFL & HEC UNIL, Lausanne, Switzerland, WCO.

Laporte B. (2011), "Risk management systems: Using data mining in developing countries' customs administrations," *World Customs Journal*, Vol. 5, No. 1, pp. 17-27.

Walsh J. T. (2003), "Vérification et audit a posteriori," in M. Keen (ed.), *Moderniser la douane: Défis et stratégies de réforme de l'administration douanière*, Washington, DC, IMF.

WCO (2003), *Guide sur la gestion des risques*, Brussels, WCO.

Widdowson D. (2005), "Managing risk in the customs context," in L. De Wulf & J. B. Sokol (eds.), *Customs Modernization Handbook*, Washington, DC, World Bank.

**FERDi**