

TC ABSTRACT

I. Basic Project Data

▪ Country/Region:	ECUADOR/CAN - Andean Group
▪ TC Name:	Evaluating Machine Learning–Driven Fraud Detection: Evidence from AquaData in EPMAPS
▪ TC Number:	EC-T1647
▪ Team Leader/Members:	PEREZ URDIALES, MARIA (INE/WSA) Team Leader; OYAMADA KROUG, JORGE RUBEN (INE/WSA) Alternate Team Leader; TOJAL RAMOS DOS SANTOS, CAROLINA (RES/RES) Alternate Team Leader; CAVALLO, EDUARDO (INE/INE); COUTO RIBEIRO BEATRIZ (INE/WSA); MOSILLO, LUDOVICA (INE/WSA); JIMENEZ ARRIOLA JUAN CARLOS (INE/WSA)
▪ Taxonomy:	Client Support
▪ Number and name of operation supported by the TC:	N/A
▪ Date of TC Abstract:	11 Sep 2026
▪ Beneficiary:	Empresa Pública Metropolitana de Agua Potable y Saneamiento de Quito (EPMAPS)
▪ Executing Agency:	INTER-AMERICAN DEVELOPMENT BANK
▪ IDB funding requested:	US\$150,000.00
▪ Local counterpart funding:	US\$0.00
▪ Disbursement period:	30 months
▪ Types of consultants:	Individuals; Firms
▪ Prepared by Unit:	INE/WSA - Water & Sanitation
▪ Unit of Disbursement Responsibility:	CAN/CEC - Country Office Ecuador
▪ TC included in Country Strategy (y/n):	No
▪ TC included in CPD (y/n):	No
▪ Alignment to IDBStrategy+ 2024-2030:	Sustainable, resilient, and inclusive infrastructure

II. Objective and Justification

- 2.1 The objective is to determine the causal impact of machine learning-driven inspection targeting, through the AquaData algorithm, on fraud detection effectiveness, inspector productivity, revenue recovery, and apparent losses at EPMAPS. Because the 2023 pilot did not generate an identifiable counterfactual, the design is prospective and experimental.
- 2.2 The main question is whether using the algorithm to prioritize inspections increases inspector productivity and fraud detection effectiveness relative to the current expert-based assignment process. Two complementary questions follow: the impact on inspector productivity, measured through inspections completed, hit rates, and recovered consumption per inspector; and the cost-effectiveness of the algorithm relative to the current process, accounting for implementation, operational, and inspection costs.
- 2.3 The knowledge gap is well defined. Utilities across the region increasingly use machine learning to detect anomalies in consumption, flow, and pressure data that may signal fraud (Lee and Kim, 2023; Romero-Ben et al., 2023), and the problem is not unique to water: electricity losses average roughly 17% of grid input in the region (Jimenez Mori et al., 2014). However, this literature remains largely descriptive and relies on retrospective comparisons that cannot separate the effect of the algorithm

from selection bias, leaving utilities without credible evidence on the gains they can expect.

- 2.4 The theory of change is direct: using existing inspection staff, the resources to operate the algorithm, and historical billing data, the project designs a data collection protocol, institutionalizes the algorithm within the commercial losses unit, and trains inspectors. The output is an operational algorithm that prioritizes connections with the highest predicted probability of fraud, which should raise the share of inspections confirming fraud and the number of inspections per day and, over time, reduce apparent losses and improve revenue recovery, fine collection, and financial sustainability.
- 2.5 The methodology is a prospective RCT randomized at the operational sector level. Of the 110 sectors (some 720,000 active connections), eligible sectors will be randomly assigned to treatment, where inspection priority is set by AquaData, or control, where the existing process continues. The sample is restricted to active residential connections (89% of contracts and 80% of consumption) and excludes remote or unsafe sectors under an ex-ante geographic criterion. Sector-level randomization mitigates spillovers between neighboring connections. The intervention is the use of the algorithm to prioritize cases, not the inspection itself, which is the field validation mechanism. Impact will be estimated by intention to treat using OLS, with standard errors clustered at the sector level and baseline controls, verifying balance between groups.
- 2.6 Outcomes are measured with EPMAPS administrative data: (a) the share of inspections confirming an infraction; (b) inspections per inspector per working day; (c) revenue recovered through billing adjustments and fines; (d) effective collection of issued fraud notifications, against a baseline of 30-40%; and (e) the apparent loss index at the operational unit level.
- 2.7 Fund resources will finance the design of the data collection protocol, the digitization of inspection records, the implementation and monitoring of the randomized rollout, the analysis, and dissemination.
- 2.8 This TC is consistent with the IDB Group's Institutional Strategy: Transformation for Greater Impact and Scale (CA-631). By recovering revenue lost to commercial fraud and improving the productivity of inspection teams, it strengthens the financial sustainability of water utilities, supporting the objective of bolstering sustainable growth, and contributes to reducing poverty and inequality, since more sustainable utilities are better positioned to expand coverage and improve service quality in underserved areas. It also responds to the operational focus areas of (i) sustainable, resilient and inclusive infrastructure, and (ii) productive development and innovation through the private sector, since AquaData leverages artificial intelligence and machine learning to modernize utility operations. It further reinforces the emphasis of the Strategy on knowledge as a comparative advantage, by generating rigorous causal evidence on AI-driven fraud detection to inform replication across the region. The TC also aligns with the IDB Group Country Strategy with Ecuador 2022-2025 (GN-3103-1), whose results matrix prioritizes improved infrastructure for water access (outcome 7.4, indicator 7.4.1, baseline 67.78% in 2019) and identifies digitalization as a crosscutting axis. It is further anchored in the Water and Sanitation Sector Framework Document, particularly its action items on efficient and sustainable service delivery and on the institutional strengthening of utilities.

III. Description of Activities and Outputs

- 3.1 **Component I: Experimental design and data collection protocol.** Finances the activities needed to put the experiment in the field: design of the data collection protocol, digitization of the currently paper-based inspection records into a database that feeds results back into the model, and finalization of the experimental design,

including randomization of the 110 operational sectors, power calculations, and outcome definition. It also covers training of inspection teams, baseline consolidation, and support to formalize data-sharing agreements with EPMAPS.

- 3.2 **Component II: Experimental implementation, monitoring and analysis** . Finances the human and technical resources needed to run and analyze the experiment. It covers operation and maintenance of the AquaData algorithm over the twelve-month implementation, including monthly generation of prioritized inspection lists in treatment sectors, and monitoring of compliance with the randomization protocol. It also funds a data scientist and an economist to process the inspection, billing, fines, and water balance data, estimate impacts, and assess cost-effectiveness.
- 3.3 **Component III: Dissemination** . This component finances editing services, the production of a working paper and a policy brief, and conference travel to support the dissemination of the knowledge generated by the evaluation, both within EPMAPS and among water utilities in the region considering the adoption of similar tools.

IV. Budget

Indicative Budget

Activity/Component	IDB/Fund Funding	Counterpart Funding	Total Funding
Experimental design and data collection protocol	US\$75,000.00	US\$0.00	US\$75,000.00
Experimental implementation, monitoring and analysis	US\$60,000.00	US\$0.00	US\$60,000.00
Dissemination	US\$15,000.00	US\$0.00	US\$15,000.00
Total	US\$150,000.00	US\$0.00	US\$150,000.00

V. Executing Agency and Execution Structure

- 5.1 The execution will be carried out by the Inter-American Development Bank through the Water and Sanitation Division (INE/WSA), under the supervision of the Team Leader and with the support of the Country Office in Ecuador (CAN/CEC). The Bank will be responsible not only for contracting, but also for monitoring, follow-up, and approval of the respective deliverables.
- 5.2 This scheme is justified by the level of inter-institutional coordination and technical knowledge required for the TC, and by the need to protect the integrity of the randomization, which requires that the assignment of sectors to treatment and control be managed independently of the utility's operational decisions.

VI. Project Risks and Issues

- 6.1 No major impacts are identified in the execution of this TC, since it involves consultancies, studies, and administrative expenses. Three operational risks are nonetheless relevant for an experimental design of this type. First, there is a risk of non-compliance with the randomization, if inspection teams deviate from the assigned prioritization method in treatment or control sectors. To mitigate this risk, the protocol will be agreed in writing with EPMAPS Commercial Management before implementation, inspectors and supervisors will be trained on the assignment rules, compliance indicators will be monitored monthly, and the impact will be estimated by intention to treat, which remains valid under imperfect compliance. Second, there is a risk of delays or gaps in the digitization of inspection records, which are currently paper-based and constitute the main source of outcome data. To mitigate this risk, digitization is financed as a distinct activity under Component I and will be completed and validated before the experimental phase begins, and data quality will

be monitored continuously throughout implementation. Third, there is a risk of delays associated with staff turnover at the utility or with the formalization of data-sharing agreements. To mitigate this risk, the Bank will lead the implementation of activities, ensure effective coordination and communication among the parties, and secure institutional rather than individual commitments from EPMAPS at the outset of execution.

VII. Environmental and Social Aspects

- 7.1 This TC does not have applicable requirements of the Bank's Environmental and Social Policy Framework (ESPF).