Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Conscious infrastructure for fairness aware detection of household water insecurity in Hyderabad informal settlements

**Anil Kumar Palakodeti**

anilkumarpalakodeti@gmail.com

**Research Article**

**Keywords:** household water insecurity, urban water governance, informal settlements, machine learning, data justice, algorithmic fairness, Hyderabad, urban informality, decision support systems

**Posted Date:** July 6th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-9707559/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations:** No competing interests reported.

# Conscious infrastructure for fairness aware detection of household water insecurity in Hyderabad informal settlements

Corresponding author: Dr Anil Kumar Palakodeti; Email: Anilkumarpalakodeti@gmail.com

Independent Consultant, Urban Water Governance and Climate Resilience, Houston, Texas, USA

## Abstract

Artificial intelligence (AI) and machine learning are increasingly used in water management for demand forecasting, water-quality prediction, leakage detection, pressure optimisation and infrastructure decision support. Much of this literature remains organised around the operational perspective of the utility: networks, assets, flows, pressures, source categories and monitoring data. This orientation is analytically incomplete for informal settlements, where water insecurity is often experienced less as the absence of infrastructure than as the household burden of making unreliable or unsafe services usable. This paper argues that AI for urban water governance cannot be fairness-aware unless it is built from the empirical structure of household water insecurity rather than infrastructure visibility alone.

The paper uses raw survey and water-quality data from Hyderabad, India, to operationalise Conscious Infrastructure Neural Systems (CINS) as a decision-support architecture that integrates household-burden data into predictive models and links outputs to specified institutional response protocols. The term "neural" is used to describe a linked decision architecture connecting household data, predictive classification, subgroup fairness audits and institutional response protocols; it does not imply that the empirical model estimated in this paper is a neural network. The analysis uses 1,700 household records across 34 informal settlements and 34 water-quality sample records. It compares a utility-visible logistic model with a household-burden model to test whether adding household-level variables improves the identification of households whose work schedules are disrupted by drinking-water supply or availability timing. This outcome is used as a proxy for binding household water insecurity because it captures direct economic and temporal consequences rather than perception alone.

The household-burden model outperformed the utility-visible model on the held-out test set. Accuracy increased from 0.669 to 0.731, recall from 0.636 to 0.714, F1-score from 0.513 to 0.593 and AUC from 0.731 to 0.818. The overall false-negative rate decreased from 0.364 to 0.286. Subgroup analysis shows that false negatives declined among the lowest-income quartile and multi-source households. The tanker-dependent subgroup showed a small deterioration, plausibly because tanker use was already a strong predictor in the utility-visible model. These results indicate that fairness in infrastructure AI is not merely a post hoc correction to model outputs; it is shaped by data inclusion, outcome definition and problem formulation.

The paper introduces the concept of the infrastructural visibility gap: the difference between the burdens households experience because of fragmented or unsafe water access and the burdens public decision systems recognise as actionable evidence. The paper formalises this gap through a CINS framework that links household-burden variables, predictive classification, subgroup false-negative audits and institutional response protocols. The findings show that fairness-aware urban water AI must be designed around household-burden detection rather than infrastructure visibility, with direct implications for how utilities prioritise interventions, allocate resources and evaluate service adequacy in informal settlements.

**Keywords**: household water insecurity; urban water governance; informal settlements; machine learning; data justice; algorithmic fairness; Hyderabad; urban informality; decision support systems

# 1. Introduction

Urban water policy frequently treats infrastructure provision as the primary indicator of progress. The operational logic is familiar: expand networks, increase treatment capacity, provide connections, improve metering, reduce non-revenue water and optimise distribution. These interventions are necessary, but they do not resolve the analytical problem at the centre of this paper: a household can be counted as connected and still remain water insecure. It may have a tap but still depend on tankers. It may be served by a municipal network but receive water too infrequently or unpredictably to organise work, care and storage safely. It may be formally visible to the utility while the costs it bears remain invisible to the decision system.

The distinction between infrastructure access and lived water security is well established in household water insecurity research. Jepson et al. (2017) argue that water insecurity must be understood through lived experience, not only through source classification or infrastructure metrics. Young et al. (2019) show that household water insecurity includes adequacy, reliability, accessibility, safety, worry and disruption. For informal settlements, this distinction is especially important because water access is rarely a single service relationship. It is an assemblage of household taps, public taps, tankers, groundwater sources, bottled water, storage practices, payments, waiting time, political mediation and household labour.

The expansion of smart-water systems and AI-based decision support creates a new version of this problem. If AI systems are trained primarily on variables that utilities already collect, such as source type, supply schedules, infrastructure condition, pressure, billing records and water-quality samples, they may optimise what is institutionally visible while missing households whose insecurity is expressed through time loss, income pressure, illness, alternative-source dependence and coping expenditure. The risk is therefore not only algorithmic bias in the narrow technical sense. It is problem-formulation bias: the model is asked to predict network-visible failure rather than household-experienced burden.

This paper examines that risk through an empirical case from Hyderabad, India. It uses a raw Excel dataset containing 1,700 household survey records across 34 informal settlements and 34 water-quality sample records. The paper compares two predictive approaches. The first is a utility-visible model built from variables closer to conventional infrastructure knowledge. The second is a household-burden model that adds income, expenditure, quality and quantity ratings, fetch burden, illness and treatment indicators. The purpose is not to claim that one model is universally superior. It is to test whether household-burden data change who becomes visible as high risk.

The paper operationalises Conscious Infrastructure Neural Systems (CINS) in a deliberately empirical manner. CINS is defined as a decision-support architecture that integrates household-burden data into predictive models and links model outputs to institutional response protocols. The term “conscious” is not used to imply sentient infrastructure. The term “neural” is used here to describe a linked decision architecture connecting household data, predictive classification, subgroup fairness audits and institutional response protocols. It does not imply that the empirical model estimated in this paper is a neural network. It identifies a design requirement: an infrastructure decision system should be capable of recognising household burdens that conventional infrastructure metrics tend to erase. In this paper, that design requirement is operationalised through a model comparison in which the reduction of false negatives functions as an empirical indicator of reduced infrastructural invisibility.

The central argument is that fairness-aware urban water AI should be evaluated by its capacity to reduce an infrastructural visibility gap. This gap emerges when households experience labour, income, health and coping burdens because of water insecurity, but public decision systems recognise only the more legible infrastructure variables. In this sense, machine learning is treated not only as a predictive tool but also as a diagnostic instrument for institutional visibility: what is missed when household burden is excluded, and who is misclassified as low risk as a result?

The paper addresses five research questions:

1. How does household water access in Hyderabad's informal settlements differ from formal infrastructure access?
2. What empirical burdens are associated with fragmented water access, including quality risk, time loss, income pressure, health burden and coping costs?
3. Which households are most likely to be missed by a utility-visible predictive model?
4. Does adding household-burden variables improve recall and reduce false negatives among vulnerable subgroups?
5. How can predicted risks be linked to institutional response protocols rather than remain passive dashboard outputs?

The contribution is threefold. First, the paper provides an empirical analysis of source fragmentation and household burden in 34 informal settlements. Second, it develops the infrastructural visibility gap as a conceptual bridge between household water insecurity, data justice and urban infrastructure governance. Third, it proposes and tests a fairness-aware modelling approach in which subgroup false negatives, rather than average accuracy alone, become a central criterion for evaluating AI-enabled water governance.

# 2. Literature Review

## 2.1 From infrastructure access to household water insecurity

Global water monitoring has historically relied on source type, infrastructure coverage and service ladders. These measures are indispensable for tracking broad progress, but they can understate insecurity when households experience intermittent supply, poor timing, unsafe storage, unaffordable alternatives or fear of running out of water. Household water insecurity research addresses this gap by treating water insecurity as a lived and multidimensional condition rather than a binary infrastructural status.

Jepson et al. (2017) call for water insecurity metrics that capture cross-disciplinary dimensions of lived experience. Young et al. (2019) developed and validated the Household Water InSecurity Experiences scale to measure experiences such as insufficient water, interruption to daily activities, worry, shame and inability to meet domestic needs. Stoler et al. (2021) further show that household-level metrics can reveal forms of insecurity that conventional monitoring misses. This literature is important for AI because it implies that source type alone is a weak input for fairness-aware prediction. A model that treats formal access as security may under-detect households that are formally connected but functionally burdened.

The Hyderabad dataset speaks directly to this argument. It allows source use, quality and quantity perceptions, income and expenditure, illness, fetch burden and work disruption to be examined together. It therefore permits water insecurity to be analysed not simply as an infrastructure condition, but as an interaction between service arrangements and household welfare consequences.

## 2.2 Urban informality, differentiated infrastructure and everyday water politics

Urban infrastructure is neither evenly distributed nor socially neutral. The literature on splintered urbanism and urban informality shows that infrastructure systems produce unequal service geographies and differentiated forms of urban citizenship (Graham and Marvin, 2001). In many low-income settlements, residents are connected through partial, intermittent or negotiated infrastructures rather than through universal and reliable service systems.

Bakker et al. (2008) argue that urban water supply failures must be understood through institutional arrangements and governance practices, not only through technical under-provision. Anand's (2011) work on Mumbai shows that water pressure is simultaneously a hydraulic condition and a political relation. Ranganathan (2014) demonstrates that payments for infrastructure are entangled with claims to recognition and citizenship. These studies show that water infrastructure is a socio-political system: it determines who is recognised, who waits, who pays, who improvises and who becomes administratively legible.

This paper extends that literature into the domain of AI decision support. If urban infrastructure already differentiates residents by visibility, entitlement and service reliability, then AI systems built on infrastructure data may reproduce these differentiations. The analytical question is therefore not only whether machine learning can predict water risk, but whether it can detect risk in households and settlements that are weakly visible to the infrastructure system.

### 2.3 Smart water, machine learning and utility-centred optimisation

Machine learning has become increasingly prominent in water research, particularly in water-quality evaluation and prediction (Zhu et al., 2022). It is also used in water systems for demand forecasting, leakage detection, pressure management, pipe-failure prediction, water-quality monitoring and regulatory compliance. These applications are valuable, but they are often built around the operational perspective of utilities. They ask how networks can be made more efficient, how assets can be maintained, how supplies can be forecast and how contaminants can be detected. They less often ask how service failures translate into household labour, income loss, health burden and informal coping.

This distinction matters because predictive systems inherit the priorities embedded in their data. Utility datasets often contain flows, sources, assets, pressure zones, billing records and maintenance logs. They rarely contain household time loss, work disruption, illness expenditure, water-fetching burdens, distrust of source quality or the need to combine multiple sources. If these variables are absent, the model cannot learn them. It can only optimise a partial representation of the water system.

This paper therefore distinguishes operational AI from governance AI. Operational AI improves network functioning. Governance AI must also evaluate whether the benefits and burdens of water service are distributed fairly. This reframes fairness in infrastructure AI not as a post hoc adjustment to model outputs, but as a property of data inclusion and problem formulation.

### 2.4 Data justice, algorithmic fairness and public-service misclassification

Data justice scholarship provides a conceptual basis for examining why household burden can disappear from technical systems. Taylor (2017) argues that data justice concerns how people are made visible, represented and treated through data. Kitchin (2014) warns that data-driven urbanism can privilege what is machine-readable while marginalising social complexity. In public infrastructure, what is easiest to measure can become what is easiest to govern.

Algorithmic fairness research adds an important warning: misclassification harms are not evenly distributed. Barocas and Selbst (2016) show that data mining can generate disparate impacts when training data inherit past exclusions or measurement gaps. Selbst et al. (2019) caution that fairness cannot be solved inside the model alone if the surrounding socio-technical system remains unjust. Public-service examples reinforce this point. In health-care risk prediction, Obermeyer et al. (2019) show how apparently neutral proxy outcomes can misallocate care. In child-welfare screening, Chouldechova et al. (2018) show how algorithm-assisted decisions require attention to error distribution and institutional context. Eubanks (2018) similarly warns that automated welfare systems can intensify administrative exclusion when social vulnerability is reduced to incomplete data categories.

The water-governance implication is direct. A false negative is not merely a statistical error. It is a household that experiences water-related burden but is classified as low risk. If such households are low-income, multi-source dependent, tanker-dependent or located in microbiologically exposed settlements, the model can become a mechanism for under-prioritisation. The literature therefore suggests a concrete evaluation criterion for CINS: it should reduce false negatives among vulnerable groups, not merely improve overall accuracy.

### 2.5 Research gap

These three bodies of scholarship identify closely related problems, but they have rarely been brought into direct analytical conversation. Household water insecurity research demonstrates that water insecurity is not

reducible to infrastructure access; it is experienced through adequacy, reliability, safety, affordability, time burden and disruption to daily life. Urban infrastructure scholarship shows that service provision is socially and politically differentiated, particularly in informal settlements where residents often depend on partial, negotiated and unevenly reliable systems. Algorithmic fairness scholarship, in turn, demonstrates that public-sector models can reproduce inequality when vulnerable groups are poorly represented in data, when proxy outcomes obscure lived harm, or when model objectives are defined around administratively visible variables rather than social consequences. The unresolved gap is therefore both empirical and methodological: there is limited evidence on whether incorporating household-burden variables into AI-enabled water-governance models improves the detection of high-risk households and reduces subgroup-specific false negatives. This paper addresses that gap through a reproducible comparison between a utility-visible model and a household-burden model, using household survey and water-quality data from informal settlements in Hyderabad.

## 3. Analytical Framework: The Infrastructural Visibility Gap

This paper develops the concept of the infrastructural visibility gap as its central analytical framework. The term refers to the disjuncture between, on the one hand, the burdens households experience because of inadequate, unsafe, unreliable, or unaffordable water access and, on the other hand, the forms of evidence that infrastructure decision systems recognise, measure, and translate into action. The concept is not intended simply to denote missing data. Rather, it identifies a governance problem that becomes algorithmic when AI-enabled decision systems are trained on partial representations of water risk. When household burdens are absent from the data architecture, they are also likely to be absent from prediction, prioritisation, and institutional response.

The infrastructural visibility gap has three interrelated dimensions: measurement, valuation, and accountability.

### 3.1 Measurement gap

The measurement gap arises when agency or utility datasets capture infrastructure status but fail to capture household experience. A settlement may appear in official records as connected to a water network, while residents continue to rely on tankers, public taps, groundwater sources, or bottled water. Similarly, utility data may record the presence of taps or supply schedules without capturing work disruption, waiting time, storage dependence, medical costs, or distrust of water quality. In predictive modelling, this gap becomes a feature-selection problem: models cannot detect burdens that were not measured, encoded, or linked to the analytical dataset. As a result, households that are formally visible as service recipients may remain analytically invisible as burdened users.

### 3.2 Valuation gap

The valuation gap occurs when household burdens are known or observable but are not treated as decision-relevant. Productivity loss, expenditure pressure, medical costs, preventive coping, and reliance on alternative water sources may be experienced by households, yet excluded from project prioritisation, feasibility assessment, tariff design, or resource allocation. In AI terms, these variables may exist in survey or field data but remain outside the model objective, the risk definition, or the decision rule. The consequence is that household suffering is acknowledged descriptively but not operationalised institutionally. A fairness-aware decision system must therefore do more than collect household-burden data; it must assign those burdens analytical weight in prediction and prioritisation.

### 3.3 Accountability gap

The accountability gap emerges when identified or predicted risks do not trigger a corresponding institutional response. A dashboard may classify a settlement or household as high risk, but if no protocol requires retesting, chlorination review, supply correction, affordability assessment, grievance escalation, or intervention

prioritisation, prediction remains informational rather than actionable. This distinction is central to CINS. A decision-support system cannot be considered fairness-aware merely because it identifies vulnerable households more accurately. Its outputs must be connected to response obligations that specify who acts, within what timeframe, using what corrective mechanism, and with what form of follow-up. CINS therefore requires not only model outputs, but also enforceable response protocols that convert algorithmic visibility into institutional accountability.

### 3.4 Operationalising the infrastructural visibility gap

The infrastructural visibility gap can be operationalised as the difference between household-experienced burden and the extent to which that burden is represented in utility-visible data. In this paper, household burden is represented through variables such as source fragmentation, work-schedule disruption, income and expenditure pressure, perceived quality and quantity concerns, illness-related indicators, fetch burden and reliance on alternative water sources. Utility visibility is represented through variables closer to conventional infrastructure knowledge, including source type, supply timing, storage duration, settlement identifier and linked water-quality risk. The expressions below are conceptual and diagnostic rather than claims that household burden and utility visibility are measured on a single common scale.

At the household level, let $B_i$denote the burden experienced by household $i$, and let $V_i$denote the portion of that burden captured by utility-visible variables. The infrastructural visibility gap can be expressed conceptually as:

$$IVG_i = B_i - V_i$$

where $IVG_i$is the unrecognised or under-detected burden experienced by household $i$.

In empirical terms, the paper does not estimate $IVG_i$as a single observed variable. Instead, it tests for the visibility gap through model comparison. If a utility-visible model misses households that are correctly identified when household-burden variables are added, then the difference in false negatives provides empirical evidence of infrastructural invisibility.

The modelling outcome is:

$$Y_i = \begin{cases} 1, & \text{if household } i \text{ reports work-schedule disruption due to drinking-water supply or availability timing} \\ 0, & \text{otherwise} \end{cases}$$

The utility-visible model is specified as:

$$Pr(Y_i = 1) = f(U_i)$$

where $U_i$includes utility-visible or infrastructure-adjacent variables such as source use, supply timing, storage duration, settlement identifier and linked water-quality risk.

The household-burden model is specified as:

$$Pr(Y_i = 1) = f(U_i, H_i)$$

where $H_i$includes household-burden variables such as income, expenditure pressure, income margin, perceived quality and quantity ratings, illness and treatment indicators, fetch burden and other household-reported welfare indicators.

The empirical test is therefore not whether AI can predict water insecurity in the abstract, but whether the inclusion of $H_i$improves the detection of households whose burden is missed when only $U_i$is used. The core visibility test can be expressed as:

$$\Delta FNR = FNR_{utility} - FNR_{burden}$$

where $FNR_{utility}$ is the false-negative rate of the utility-visible model and $FNR_{burden}$ is the false-negative rate of the household-burden model. A positive $\Delta FNR$ indicates that the household-burden model reduces under-detection.

For subgroup fairness, the same logic is applied to vulnerable subgroup $g$:

$$\Delta FNR_g = FNR_{utility,g} - FNR_{burden,g}$$

where $g$ may refer to the lowest-income quartile, tanker-dependent households, multi-source households or households located in microbiological-risk settlements. This formulation links the infrastructural visibility gap to the distribution of model errors. CINS is therefore evaluated not only by overall predictive performance, but by whether it reduces false negatives among households for whom misclassification carries higher welfare consequences.

Finally, the governance requirement of CINS can be expressed as:

$$R_i = \phi(\hat{Y}_i, G_i)$$

where $\hat{Y}_i$ is the predicted household risk and $G_i$ represents the relevant institutional response protocol. The purpose of this formulation is to emphasise that prediction alone is insufficient. A household classified as high risk should trigger a response pathway, such as water-quality retesting, supply-schedule review, affordability assessment, grievance escalation or prioritised intervention. In this sense, CINS links prediction to institutional accountability.

The framework illustrates how household-experienced water insecurity is translated into a data architecture, model comparison, subgroup false-negative audit and institutional response protocol. It treats prediction as incomplete unless linked to response obligations.

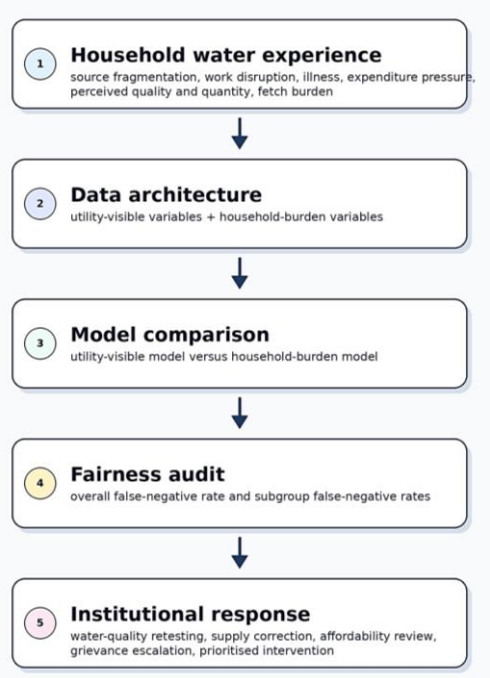



***Figure 1: CINS framework linking household burden, predictive classification, fairness audit and institutional response***

### 3.5 Propositions tested in the analysis

The empirical analysis is structured around five propositions that link the conceptual framework to the model comparison and subgroup fairness tests.

1. **Formal access is not equivalent to secure access.**
   The presence of an individual household tap does not necessarily indicate water security. Households may remain water insecure when formal access is accompanied by intermittent supply, unreliable timing, poor perceived quality, storage dependence, or continued reliance on supplementary sources such as tankers, public taps, groundwater, or bottled water.
2. **Multi-source dependence is an indicator of service fragmentation.**
   Reliance on multiple drinking-water sources is treated not simply as household choice or adaptive flexibility, but as a potential signal of incomplete primary service. Where households combine individual taps with tankers, public taps, groundwater, or bottled water, this may indicate that no single source reliably satisfies household requirements for quantity, quality, timing, affordability, and trust. The analysis therefore tests whether multi-source dependence is associated with reported work-schedule disruption.
3. **Household-burden variables improve the detection of high-risk households.**
   A model that incorporates household-burden variables—income, expenditure pressure, perceived water quality and quantity, illness, fetch burden, treatment practices, and other welfare-relevant indicators—is expected to identify disrupted households more effectively than a model based only on utility-visible variables. This proposition tests whether household experience adds predictive value beyond infrastructure and source-access indicators.
4. **Fairness must be assessed through the distribution of false negatives across vulnerable subgroups.**
   In water-service decision-making, the most consequential modelling error is not only misclassification in general, but the false negative: a household experiencing burden but classified as low risk. The analysis therefore evaluates whether model errors are unevenly distributed among income-vulnerable households, multi-source households, tanker-dependent households, and households located in microbiological-risk settlements. This shifts fairness evaluation from aggregate accuracy to subgroup-specific under-detection.
5. **CINS requires institutional response, not prediction alone.**
   A model that identifies water-risk households but does not trigger corrective action produces visibility without accountability. For CINS to function as a governance architecture, predictions must be linked to response protocols such as retesting, chlorination review, supply-schedule correction, affordability assessment, grievance escalation, or prioritised intervention. The analysis therefore treats model performance as only one component of fairness-aware water governance; institutional follow-through is the second.

## 4. Materials and Methods

### 4.1 Dataset and analytical sample

The empirical analysis uses the Excel workbook associated with the Hyderabad field survey as the final and authoritative dataset. The workbook contains 34 settlement-level household survey sheets and one water-quality data sheet. The final analytical sample consists of 1,700 household records, with 50 records per settlement, and 34 water-quality sample records. Earlier figures used in concept-note drafts are not used in this paper; all descriptive statistics and model results are calculated from the Excel workbook.

The dataset originates from doctoral research on drinking-water access, quality, quantity and decision-making in Hyderabad's informal settlements (Palakodeti, 2021). The present paper uses those data for a secondary, reproducible modelling analysis.

***Table 1. Final analytical sample***

| Dataset component | Final sample used in this paper |
|---|---|

| | |
|---|---|
| Household survey records | 1,700 |
| Slum-area sheets / settlement strata | 34 |
| Household records per settlement | 50 |
| Water-quality sample records | 34 |

*Note. All counts are calculated from the Excel workbook used for this analysis.*

## 4.2 Variable construction and data preparation

The household-level dataset was constructed by merging the 34 settlement-level sheets into a single analytical file with one row per household. Each household record was assigned a settlement identifier corresponding to its original sheet. This created a final household dataset of 1,700 records, with 50 records per settlement.

Source-use indicators were constructed as binary variables. Separate indicators were created for individual tap use, public tap use, groundwater-source use, tanker use and bottled-water use. Because households could report more than one drinking-water source, these categories were not treated as mutually exclusive. A multi-source dependence variable was created by summing the source indicators for each household and coding households as multi-source dependent when they reported more than one source.

Household-burden variables were constructed from available survey fields capturing income, expenditure, water-related expenditure, quality and quantity ratings, fetch burden, illness indicators, treatment practices and work-schedule disruption. Where income and expenditure were recorded in categorical bands, midpoint values were used to approximate continuous measures. These midpoint approximations were used to construct income, expenditure, expenditure ratio and income-margin variables. Because such approximations introduce measurement uncertainty, the resulting estimates are treated as modelling inputs rather than precise household-income measurements.

The water-quality sheet was cleaned separately and used to construct microbiological-risk indicators. The cleaned water-quality records were linked to household records where settlement-level matching was possible. This linkage was used to create contextual risk variables for households located in settlements where the available water-quality sample indicated microbiological risk. The linked variable should therefore be interpreted as a settlement-level exposure signal rather than a direct measure of the water consumed by each household.

The modelling dataset was then divided into two feature sets. The utility-visible feature set included variables closer to conventional infrastructure or utility knowledge, such as source use, supply timing, storage duration, settlement identifier and linked water-quality risk. The household-burden feature set added variables capturing household welfare consequences, including income and expenditure indicators, quality and quantity ratings, illness and treatment indicators, fetch burden and other household-reported burden variables.

Missing or unclear values were handled conservatively. Variables with insufficient consistency for modelling were excluded from the feature set. Binary indicators were coded only where the underlying response could be interpreted clearly. The purpose of data preparation was not to maximise model complexity, but to create a reproducible and interpretable dataset suitable for comparing the predictive value of utility-visible and household-burden variables.

## 4.3 Water-quality risk classification

The water-quality sheet was cleaned as a separate dataset with 34 sample records. The samples included municipal water, groundwater and tank-water records. Microbiological risk was coded as one if total coliforms were detected above zero or if thermotolerant coliforms or E. coli were recorded as present. This coding is intentionally conservative: it treats microbiological detection as a risk signal suitable for prioritisation, while recognising that the samples are settlement/source observations rather than household-level stored-water tests. The selected indicators are consistent with rapid drinking-water quality assessment approaches that emphasise microbiological risk as a critical public-health signal (Howard et al., 2003).

Where matching by settlement was possible, the microbiological-risk flag was joined to the household records. This linkage does not imply that every household in a matched settlement consumed contaminated water. Instead, it identifies households located in settlements where the sampled water-quality environment indicated microbiological risk. The model therefore uses water-quality data as a contextual exposure variable rather than as a household-level diagnosis.

Across the full household dataset, 900 of 1,700 households were located in settlements with linked water-quality sample records. Of these, 750 households were located in settlements where the linked water-quality sample indicated microbiological risk. In the held-out test set, 257 households were located in settlements with linked water-quality samples, and 224 were located in settlements where the linked sample indicated microbiological risk. These counts clarify that the water-quality variables are not household-level exposure measurements; they are settlement-linked contextual risk indicators used to test whether microbiological-risk environments are associated with model under-detection.

## 4.4 Outcome variable and justification

The primary outcome variable is whether the household reported that work schedules were affected by drinking-water supply or availability timing. This variable is used as a proxy for binding household water insecurity because it captures direct economic and temporal consequences, which are less subjective than perception-based measures alone. It indicates that water access is not merely inconvenient but disruptive enough to interfere with income-generating routines or household labour allocation.

The outcome has limitations. It captures one consequential dimension of water insecurity, not the full spectrum of welfare impacts. It does not directly measure all illness, gendered labour, affordability or psychological stress. However, for the present modelling exercise it is analytically useful because it connects water access to a concrete household-level burden and allows false-negative analysis: which households report disruption but are missed by different model specifications?

## 4.5 Model design

Two logistic-regression models were estimated. Logistic regression was selected as the primary model because it provides transparent classification results and interpretable coefficients. A random-forest feature-importance analysis was used as a supplementary interpretability check, not as the primary fairness model. Feature importance was calculated using the impurity-based importance scores from the fitted random-forest classifier. These scores identify variables that contributed to the model's classification structure but should not be interpreted as causal effects.

Model A is a utility-visible model. It represents the type of prediction that could be built from variables closer to conventional utility knowledge: source use, supply frequency, storage duration, timing indicators, distance or fetch variables where available, settlement identifiers and linked water-quality risk. Model B is a household-burden model. It includes the utility-visible variables and adds household-burden and vulnerability variables such as income midpoint, expenditure estimates, expenditure ratio, income margin, monthly water expenditure, quality and quantity ratings, illness and treatment variables, and other household-reported burden indicators available in the workbook.

The two models were constructed from different feature sets, as shown in Table 2. The utility-visible model approximates the kind of prediction that could be produced from infrastructure-adjacent data. The household-burden model retains those variables and adds household-level welfare and burden indicators.

***Table 2. Feature sets used in model comparison***

| Feature group | Utility-visible model | Household-burden model |
|---|---|---|
| Source variables | Individual tap, public tap, groundwater source, tanker, bottled water, source count, multi-source dependence | Same as utility-visible model |

| | | |
|---|---|---|
| Service variables | Supply interval, supply hours, timing regularity, bucket quantity, storage duration, fetch/distance variables where available | Same as utility-visible model |
| Water-quality variables | Linked microbiological-risk indicator, E. coli indicator, thermotolerant-coliform indicator, total-coliform-risk indicator | Same as utility-visible model |
| Settlement/context variables | Settlement identifier and available contextual controls | Same as utility-visible model, with additional household-context variables where available |
| Economic variables | Not included | Income midpoint, estimated expenditure, water expenditure, expenditure ratio, income margin |
| Perception/rating variables | Not included | Average source rating, minimum source rating, low quality/quantity rating, issue count |
| Health/treatment variables | Not included | Illness indicators, diarrhoea, doctor fee, medicine fee, medical expenditure, treatment practices |
| Household-burden variables | Not included | Fetch burden, treatment count, household structure, fetch burden, household structure and welfare-related indicators, with outcome-proximate work-disruption variables excluded |

*Note. Variables that directly reproduce the outcome variable were excluded from the feature set to avoid leakage. The model comparison is designed to test whether household-burden variables add predictive value beyond utility-visible variables.*

The comparison is deliberately structured as an argument. If Model B improves recall and reduces false negatives among vulnerable subgroups, then household-burden data are not merely descriptive additions; they are decision-relevant variables. If Model B fails to improve subgroup detection, then CINS would require either better burden variables or a different modelling strategy. The model comparison therefore tests the empirical value of household-burden visibility.

### 4.6 Model evaluation and fairness tests

The data were split into a 70% training set and a 30% held-out test set using stratification on the outcome variable and a fixed random seed of 42. Model performance was evaluated using accuracy, precision, recall, F1-score, AUC and false-negative rate. Five-fold cross-validation was used as a robustness check.

For selected logistic-regression predictors used in the interpretability discussion, 95% bootstrap intervals were estimated using 500 stratified bootstrap resamples of the training data. These intervals are reported for interpretability only. Because the primary model uses a scaled and class-weighted logistic specification, the intervals should be interpreted as uncertainty around the model-estimated association rather than as causal confidence intervals.

The most important metric for this paper is not average accuracy. In public-service allocation, the central fairness concern is the false negative: a household that experiences disruption but is classified as not high risk. The analysis therefore reports subgroup false-negative rates for the lowest-income quartile, tanker-dependent households, multi-source households, households in microbiological-risk settlements and households in quality-sampled settlements. This makes fairness an empirical test rather than a normative assertion.

### 4.7 Reproducibility

The analysis was designed to be reproducible from the original Excel workbook and the accompanying analysis script. The 34 settlement-level household sheets were consolidated into a single cleaned household dataset, and the water-quality sheet was processed separately before being linked to the household records where settlement-level matching was possible. All descriptive statistics, model-performance results, subgroup

false-negative-rate tables, prediction outputs and feature-importance results were generated from this cleaned analytical dataset.

Where income and expenditure variables were recorded in categorical bands, midpoint values were used to approximate continuous measures. These approximations were applied consistently across the dataset and are documented in the reproducibility script. The modelling workflow uses a fixed random seed for the train-test split and applies the same feature definitions across the utility-visible and household-burden models, ensuring that differences in model performance arise from the inclusion of household-burden variables rather than from inconsistent preprocessing.

The reproducibility package includes the cleaned household dataset, cleaned water-quality dataset, model-output tables, subgroup-fairness outputs, prediction files, feature-importance files and the analysis script. These materials can be submitted as supplementary files subject to the requirements of the target journal, anonymisation standards and data-governance considerations. Because the analysis uses household-level field-survey data, any shared dataset should exclude personally identifying information and should remain consistent with the ethical conditions under which the original data were collected.

### 4.8 Ethical and data-governance considerations

The use of household-level data in AI-enabled water governance raises ethical and data-governance concerns. The purpose of the modelling exercise in this paper is not to classify households for exclusion, surveillance or service denial. Rather, the purpose is to identify households and settlements whose water-related burdens may be under-recognised by conventional infrastructure data and to support more equitable prioritisation.

No personally identifying household information is reported in this manuscript. Model outputs are presented in aggregate form, and subgroup analysis is used to evaluate under-detection rather than to stigmatise specific communities. The paper treats high-risk classification as a trigger for support, investigation and institutional response, not as a basis for punitive targeting.

A fairness-aware water-governance system should also allow for contestability. Households and communities should have opportunities to challenge, correct or contextualise data used in decision-making. This is particularly important in informal settlements, where administrative data may be incomplete and where service histories are often shaped by negotiation, informality and uneven institutional recognition.

For this reason, CINS is conceptualised as a decision-support architecture rather than an automated decision-making system. Its outputs should inform institutional review, water-quality retesting, supply correction, affordability assessment and grievance escalation. They should not replace professional judgement, community engagement or public accountability. The ethical value of CINS depends not only on predictive performance, but on whether its use expands recognition and improves response for households experiencing water insecurity.

## 5. Results

### 5.1 Household-source fragmentation

The household records show substantial source fragmentation. Individual taps were the most common reported source, but alternative-source reliance was also high. Of the 1,700 households, 1,127 used individual taps, 676 used tankers, 448 used groundwater sources, 408 used public taps and 58 used bottled water. More than half of the households, 951 records or 55.9%, used more than one drinking-water source.

***Table 3. Household source distribution, N = 1,700***

| Source indicator | Household count | Share of households |
|---|---|---|
| Individual tap | 1,127 | 66.3% |
| Public tap | 408 | 24.0% |

| | | |
|---|---|---|
| Groundwater source | 448 | 26.4% |
| Water tanker | 676 | 39.8% |
| Bottled water | 58 | 3.4% |
| Multiple-source household | 951 | 55.9% |

*Note. Source indicators are not mutually exclusive; households may report more than one source.*

This result is analytically important because it shows that household water access cannot be represented by a single infrastructure category. Multiple-source dependence is not automatically resilience. In this context it is more plausibly a signal of incomplete primary service: households combine sources because no single source reliably satisfies quantity, timing, quality and trust requirements. For CINS, source overlap should therefore be treated as a risk feature rather than removed as noise.

## 5.2 Water-quality risk

The water-quality dataset contained 19 municipal-water samples, 13 groundwater samples and 2 tank-water samples. Microbiological risk was flagged where total coliforms were detected above zero or where thermotolerant coliforms or E. coli were recorded as present. Using this definition, 22 of 34 samples showed microbiological risk. Thermotolerant coliforms were present in 15 samples and E. coli was present in 9 samples.

***Table 4. Water-quality risk summary, N = 34 samples***

| Indicator | Count | Share of samples |
|---|---|---|
| Municipal-water samples | 19 | 55.9% |
| Groundwater samples | 13 | 38.2% |
| Tank-water samples | 2 | 5.9% |
| Microbiological-risk samples | 22 | 64.7% |
| Thermotolerant coliform present | 15 | 44.1% |
| E. coli present | 9 | 26.5% |

*Note. Microbiological-risk flag is based on total coliforms, thermotolerant coliforms or E. coli detection in the available water-quality sheet.*

The water-quality results do not support a simple distinction between formal and safe sources. They indicate that a fairness-aware decision system must incorporate source-specific and settlement-linked risk signals. The implication is not that the 34 samples provide exhaustive household-level exposure measurement, but that quality risk is sufficiently present to be treated as a contextual variable in prioritisation.

## 5.3 Model comparison

The modelling outcome was work-schedule disruption linked to drinking-water supply or availability timing. In the final analytical dataset, 468 of 1,700 households reported this burden, giving an outcome prevalence of 27.5%. The household-burden model outperformed the utility-visible model on the held-out test set. Accuracy increased from 0.669 to 0.731, recall from 0.636 to 0.714, F1-score from 0.513 to 0.593 and AUC from 0.731 to 0.818. The overall false-negative rate decreased from 0.364 to 0.286.

***Table 5. Model performance on held-out test set***

| Model | Accuracy | Precision | Recall | F1 | AUC | False-negative rate |
|---|---|---|---|---|---|---|
| Utility-visible model | 0.669 | 0.430 | 0.636 | 0.513 | 0.731 | 0.364 |
| Household-burden model | 0.731 | 0.508 | 0.714 | 0.593 | 0.818 | 0.286 |

*Note. Held-out test set: n = 510; test-positive rate = 27.5%.*

To make the classification results more transparent, Table 6 reports the confusion matrices for both models on the held-out test set. The utility-visible model identified 89 true positives and missed 51 households that

reported work-schedule disruption. The household-burden model identified 100 true positives and reduced the number of missed disrupted households to 40. It also reduced false positives from 118 to 97.

***Table 6. Confusion matrix on held-out test set***

| Model | True negatives | False positives | False negatives | True positives |
|---|---|---|---|---|
| Utility-visible model | 252 | 118 | 51 | 89 |
| Household-burden model | 273 | 97 | 40 | 100 |

*Note. True positives are households correctly classified as reporting work-schedule disruption. False negatives are households that reported work-schedule disruption but were classified as not high risk.*

Together, the performance metrics and confusion matrices show that the household-burden model improves discrimination while reducing the number of burdened households misclassified as low risk. This supports the paper's claim that household-burden variables contain decision-relevant information not captured by utility-visible variables alone.

Five-fold cross-validation produced a similar pattern. Mean recall increased from 0.613 in the utility-visible model to 0.707 in the household-burden model. Mean F1-score increased from 0.507 to 0.593 and mean AUC increased from 0.733 to 0.795.

***Table 7. Five-fold cross-validation results***

| Model | Mean accuracy | Mean recall | Mean F1 | Mean AUC |
|---|---|---|---|---|
| Utility-visible model | 0.672 | 0.613 | 0.507 | 0.733 |
| Household-burden model | 0.734 | 0.707 | 0.593 | 0.795 |

*Note. Cross-validation used stratified folds and the same feature sets as the held-out test analysis.*

## 5.4 Subgroup fairness results

The key fairness test was whether the household-burden model reduced false negatives among vulnerable subgroups. The results show improvement for most, but not all, subgroups. Among the lowest-income quartile, the false-negative rate decreased from 32.1% to 25.0%. Among multi-source households, it decreased from 40.5% to 31.6%. Among households in microbiological-risk settlements, it decreased slightly from 28.0% to 26.7%. Among households located in quality-sampled settlements, it decreased from 31.7% to 26.8%.

The exception is tanker-dependent households. For this subgroup, the false-negative rate increased from 8.8% in the utility-visible model to 12.5% in the household-burden model. This result is substantively interpretable. Tanker use is already directly visible in Model A, so it functions as a strong utility-visible signal of disruption. When additional household-burden variables are added in Model B, model weights are redistributed across income, expenditure, ratings, illness and fetch burden. This improves overall detection but slightly weakens performance for the subgroup whose main risk signal was already captured in the utility-visible model. The result cautions against assuming that adding social variables automatically improves fairness for every subgroup.

***Table 8. Subgroup false-negative-rate tests***

| Subgroup | Utility-visible FNR | Household-burden FNR | FNR reduction | Test-set subgroup N | Actual positives |
|---|---|---|---|---|---|
| Lowest-income quartile | 0.321 | 0.250 | +0.071 | 108 | 28 |
| Tanker-dependent households | 0.088 | 0.125 | -0.038 | 204 | 80 |
| Multi-source households | 0.405 | 0.316 | +0.089 | 294 | 79 |
| Microbiological-risk settlements | 0.280 | 0.267 | +0.013 | 224 | 75 |
| Quality-sampled settlements | 0.317 | 0.268 | +0.049 | 257 | 82 |

*Note. Positive FNR reduction means the household-burden model reduced false negatives relative to the utility-visible model.*

The subgroup results strengthen rather than weaken the analysis because they show that fairness is not a single global property of a model. A model can improve overall performance while producing uneven effects across groups. For urban water governance, CINS should therefore include routine subgroup audits rather than rely on aggregate performance metrics.

## 5.5 Interpretability and feature importance

Interpretability analysis supports the substantive logic of the model comparison. In the utility-visible logistic model, multi-source dependence had a positive association with work-schedule disruption, with an odds ratio of 2.19 and a 95% bootstrap interval of 1.51–3.49. Tanker use also had a positive association, with an odds ratio of 1.39 and a 95% bootstrap interval of 1.19–1.64. These results indicate that source fragmentation is not only descriptively important but also predictive of household disruption.

In the household-burden random-forest model, the highest-ranked predictors by feature-importance score included estimated monthly expenditure, supply hours, expenditure ratio, income margin, monthly water expenditure, average source rating, fetch time, tanker dependence, minimum source rating, individual tap use and timing regularity. These variables indicate that the model's classification was not driven by infrastructure access alone. Rather, economic pressure, perceived service quality, fetch burden, source dependence and supply timing jointly contributed to the model's ability to distinguish households reporting work-schedule disruption from those not reporting such disruption.

***Table 9. Selected interpretability results***

| Model/output | Variable | Result | Interpretation |
|---|---|---|---|
| Utility logistic regression | Multi-source household | OR = 2.19; 95% bootstrap interval = 1.51–3.49 | Multi-source dependence is associated with higher odds of work-schedule disruption |
| Utility logistic regression | Tanker use | OR = 1.39; 95% bootstrap interval = 1.19–1.64 | Tanker dependence is associated with higher odds of work-schedule disruption |
| Household-burden random forest | Estimated expenditure | Feature importance = 0.057 | Economic burden contributes to model discrimination |
| Household-burden random forest | Supply hours | Feature importance = 0.055 | Service timing and availability contribute to household-risk classification |
| Household-burden random forest | Expenditure ratio | Feature importance = 0.049 | Affordability pressure contributes to model discrimination |
| Household-burden random forest | Average source rating | Feature importance = 0.040 | Perceived service quality contributes to model discrimination |
| Household-burden random forest | Fetch time | Feature importance = 0.039 | Labour burden contributes to household-risk classification |

*Note. Logistic-regression odds ratios indicate the direction and magnitude of association between selected predictors and reported work-schedule disruption. Bootstrap intervals were estimated using 500 stratified resamples of the training data. Random-forest feature importance is impurity-based feature importance from the fitted random-forest classifier and is used as an interpretability supplement. These results should not be interpreted as causal effects.*

## 5.6 Robustness check and sensitivity considerations

The main robustness check was five-fold cross-validation using the same feature sets as the held-out test analysis. The cross-validation results were consistent with the primary model comparison. Mean recall increased from 0.613 in the utility-visible model to 0.707 in the household-burden model. Mean F1-score increased from 0.507 to 0.593, and mean AUC increased from 0.733 to 0.795. This indicates that the performance improvement of the household-burden model was not limited to a single held-out split.

The results should nevertheless be interpreted as a transparent baseline rather than as an optimised modelling exercise. Logistic regression was selected because it allows interpretable comparison between utility-visible

and household-burden feature sets. More complex models may improve predictive performance, but they may also reduce interpretability or produce different subgroup-fairness trade-offs. For this reason, the random-forest analysis is used only as an interpretability supplement, not as the primary fairness model.

Two sensitivity considerations are especially important for future implementation. First, the results reported in this paper use a common classification threshold for both models. In operational settings, a lower threshold may be appropriate where the policy objective is to minimise false negatives among vulnerable households. Future applications of CINS should therefore test alternative thresholds and evaluate whether subgroup false-negative rates change under different decision rules.

Second, outcome sensitivity should be examined in future work. The present analysis uses work-schedule disruption as the primary outcome because it captures a concrete economic and temporal burden. However, household water insecurity is multidimensional. A broader high-burden outcome combining work disruption, illness, water-related expenditure, quality dissatisfaction, quantity dissatisfaction and multi-source dependence may produce different model results. Future analysis should therefore test whether the household-burden model remains superior when the outcome is defined as cumulative water insecurity rather than work disruption alone.

These considerations reinforce the central interpretation of the paper. The analysis does not claim to identify a final optimal model. It demonstrates that household-burden variables improve detection under a transparent baseline specification and that subgroup false-negative rates provide a more policy-relevant fairness test than aggregate accuracy alone.

### 5.7 Risk-to-response protocol

The model outputs should not be treated as passive dashboard indicators. For governance relevance, predicted risk must trigger response protocols. The response protocol below translates CINS from a predictive tool into an institutional decision-support process.

***Table 10. CINS risk-to-response protocol***

| Predicted risk | Trigger | Response protocol |
|---|---|---|
| High microbiological risk | Total coliforms, thermotolerant coliforms or E. coli detected or predicted | Immediate retesting, chlorination review, source-level inspection and public-health advisory where relevant |
| High quantity/service risk | Source fragmentation, long supply interval, low source rating or storage stress | Supply schedule correction, tanker regulation, emergency supply review and source-reliability audit |
| High household-burden risk | Predicted work-schedule disruption or high burden | Prioritise household/settlement in intervention planning; review affordability and labour burden |
| High infrastructural visibility gap | Utility-visible model misses a household identified by the burden model | Mandatory household-burden review before project prioritisation |

*Note. The protocol is intended to connect prediction to institutional action rather than leave outputs as passive dashboard scores.*

## 6. Discussion

### 6.1 Theoretical contribution: fairness as problem formulation

The results shift the discussion of AI in water governance away from a narrow question of predictive performance. The key finding is not simply that one model has higher AUC or recall than another. The deeper finding is that the definition of the decision problem changes who becomes visible as high risk. When the model is restricted to utility-visible variables, it misses more households whose work schedules are disrupted

by water access. When household-burden variables are included, the model identifies more of these households and reduces false negatives in several vulnerable subgroups.

This reframes fairness in infrastructure AI not as a post hoc adjustment to model outputs, but as a property of data inclusion and problem formulation. Fairness begins before model training: in deciding whether household labour, expenditure pressure, illness, source ratings and multi-source dependence are treated as legitimate infrastructure evidence. The paper therefore contributes to algorithmic fairness debates by showing how a public-infrastructure model can be unfair even if technically competent, because it is built around an incomplete representation of harm.

CINS adds a governance-oriented interpretation to smart-water debates. Conventional smart-water systems often ask how to optimise infrastructure performance. CINS asks which households remain under-detected when infrastructure performance is the dominant evidence base. The paper's main contribution is therefore not the model alone, but the model comparison as a diagnostic of institutional visibility.

### 6.2 Empirical contribution: water insecurity as service fragmentation plus burden

The Hyderabad results show that household water insecurity is better understood as service fragmentation plus burden than as a simple absence of infrastructure. Individual taps were common, but tanker use, groundwater reliance, public tap use and multi-source dependence remained substantial. More than half of surveyed households used multiple sources. This undermines any binary classification of households as served or unserved.

The interpretive point is important for policy. Multi-source use can be framed positively as flexibility or adaptation, but the modelling results suggest that it is associated with disruption. In low-income informal settlements, multiple-source use often reflects the failure of a primary source to provide sufficient, reliable or trusted water. For utilities, alternative-source reliance should therefore be treated as an early warning of incomplete service, not simply as evidence that households have options.

The water-quality findings further complicate infrastructure-centered interpretations. The detection of microbiological risk in a majority of available water-quality samples indicates that source categories do not automatically map onto safety. In an AI decision system, quality exposure should therefore be modelled as a contextual risk variable and linked to response protocols such as retesting, chlorination review and public-health advisories.

### 6.3 Why false negatives matter for urban water governance

False negatives are ethically and operationally important because they represent households that experience disruption but are classified as not high risk. In public-service delivery, such households may be deprioritised, excluded from corrective interventions or made invisible in resource allocation. A model that performs well on average can still be inequitable if it misses the households for whom the consequences of misclassification are most severe.

The subgroup analysis demonstrates why aggregate accuracy is inadequate. The household-burden model improved overall recall and reduced false negatives among the lowest-income quartile and multi-source households. These are precisely the groups for whom under-detection matters: low-income households have less capacity to absorb coping costs, and multi-source households are likely to be managing incomplete service through labour and expenditure. By contrast, the tanker-dependent subgroup result shows that improvements are not automatic. Fairness must be audited across multiple vulnerability categories rather than inferred from overall performance.

### 6.4 Interpreting the tanker-dependent anomaly

The tanker-dependent result deserves careful interpretation because it prevents a simplistic claim that more household data always improve every fairness outcome. Tanker dependence was already included in the

utility-visible model and appears to be a strong signal of disruption. Adding household-burden variables improved the model's overall ability to detect high-burden households but redistributed predictive weight across many variables. This likely introduced a small trade-off for tanker-dependent households because the utility model had already captured their risk signal efficiently.

This anomaly has two implications. First, fairness-aware modelling must be iterative. If a subgroup deteriorates, the model should be recalibrated, threshold-adjusted or supplemented with subgroup-specific response rules. Second, CINS should not be treated as a single model but as a governance process that includes model comparison, subgroup auditing and institutional review. The goal is not to prove that one model is universally fair; it is to create a system that detects where unfairness remains.

### 6.5 Policy implication: household burden must become infrastructure evidence

For utilities and urban local bodies, the practical implication is clear: risk dashboards should not only show supply zones, source types, pressure values or water-quality samples. They should also show where households experience labour and income burdens because of water insecurity. A settlement with a formal source but high work disruption, multi-source dependence and microbiological risk should not be treated as adequately served.

Without incorporating household-burden indicators, smart-water investments risk systematically under-prioritising settlements that are formally connected but functionally water-insecure. This is the central policy hook of the paper. The problem is not only that poor households lack infrastructure. It is that existing infrastructure categories can make their insecurity appear solved. CINS is useful only if it corrects this representational error.

The risk-to-response protocol is therefore not an optional implementation detail. It is part of the fairness design. If high microbiological risk is detected, retesting and treatment review should follow. If high service risk is detected, supply scheduling and tanker regulation should be reviewed. If high household burden is detected, affordability and labour burden should be considered in prioritisation. If the household-burden model identifies households missed by the utility-visible model, the case should trigger an institutional visibility review.

### 6.6 Limitations and boundaries of the claims

The analysis has several limitations that define the scope of the claims advanced in this paper. First, the household dataset is cross-sectional. It supports classification, model comparison, and burden detection, but it does not support temporal forecasting. Accordingly, the paper does not claim short-term prediction of shortage cycles, five-day forecasting, or LSTM-based temporal modelling.

Second, the water-quality dataset contains 34 sample records. These records are useful for identifying contextual microbiological risk at the settlement or source level, but they should not be interpreted as exhaustive household-level exposure measurements. The linked water-quality variable therefore functions as a contextual risk indicator rather than a direct measure of the water consumed by each household.

Third, the primary outcome variable captures one specific dimension of household water insecurity: reported disruption to work schedules caused by drinking-water supply or availability timing. This is a consequential and policy-relevant indicator because it links water access to time and economic burden. However, it does not capture the full range of household water insecurity, including psychological stress, gendered labour, long-term health effects, affordability pressures, storage-related risks, or intra-household coping strategies.

Fourth, income and expenditure variables required midpoint approximations where the original survey recorded values in categorical bands. These approximations are appropriate for exploratory modelling and subgroup analysis, but they introduce measurement uncertainty. Future work using continuous income and expenditure data would allow more precise estimation of affordability and welfare-loss effects.

Fifth, logistic regression was used as the primary modelling approach because of its interpretability and suitability for a first empirical test of the CINS framework. Other modelling approaches, including gradient boosting, random forests, Bayesian models, or fairness-constrained classifiers, may produce different performance and subgroup-fairness trade-offs. The results should therefore be understood as evidence from a transparent baseline model rather than as the final or optimal modelling specification.

These limitations are not incidental weaknesses but boundaries of inference. They ensure that the paper's claims remain proportionate to the available data. The results should be interpreted as evidence that, within this Hyderabad dataset, household-burden variables improve the detection of reported work-schedule disruption and reduce false negatives for several vulnerable subgroups. They should not be interpreted as evidence that CINS has been fully deployed, operationally validated, or generalised across multiple urban contexts. Future research should test the framework using larger samples, repeated observations, richer water-quality monitoring, and institutional implementation data.

### 6.7 Future research

Future research should extend this work in four directions.

First, subsequent studies should develop a more comprehensive outcome measure for high-burden household water insecurity. The present analysis uses work-schedule disruption as a consequential and observable proxy, but future work should construct and validate a composite outcome that integrates multiple dimensions of burden, including work disruption, illness, medical expenditure, coping costs, perceived quality and quantity concerns, source fragmentation, storage dependence, and reliance on alternative water sources. Such an outcome would allow researchers to distinguish between households experiencing inconvenience and those facing cumulative welfare loss.

Second, future research should collect repeated household, service, and water-quality observations over time. The present dataset is cross-sectional and therefore supports classification rather than temporal forecasting. Longitudinal data would make it possible to examine how household burden changes across seasons, supply cycles, contamination events, heat periods, infrastructure repairs, or tanker-service interruptions. Only with such temporal depth would it be methodologically appropriate to evaluate more complex forecasting models, including recurrent neural networks or other sequence-based methods.

Third, future work should examine model calibration and decision thresholds across vulnerable subgroups. The tanker-dependent subgroup result in this paper shows that improvements in aggregate model performance do not necessarily translate into uniform fairness gains. Future studies should therefore test subgroup-specific calibration, threshold adjustment, cost-sensitive learning, and fairness-constrained optimisation to reduce false negatives among groups for whom misclassification carries higher welfare consequences.

Fourth, future research should move beyond predictive performance and study institutional uptake. The central test of CINS is not whether a model can generate risk scores, but whether those scores change public decision-making. Future studies should therefore examine whether utilities and urban local bodies respond differently when household-burden predictions are incorporated into supply scheduling, water-quality surveillance, grievance redress, affordability assessment, and investment prioritisation. This would shift the evaluation of fairness-aware water AI from model accuracy to institutional accountability and service outcomes.

## 7. Conclusion

This paper operationalised Conscious Infrastructure Neural Systems (CINS) through a reproducible, fairness-aware analysis of household water burden in Hyderabad's informal settlements. Using 1,700 household survey records and 34 water-quality sample records, the study compared a utility-visible model with a household-burden model to examine whether incorporating household-level burden variables improves the identification of households whose work schedules are disrupted by drinking-water supply or availability timing.

The findings show that household-burden variables improve overall model performance and reduce false negatives. More importantly, they show that these improvements are not evenly distributed across all subgroups. The household-burden model reduced false negatives among the lowest-income quartile and multi-source households but produced more uneven results for other groups. This pattern is analytically significant. It demonstrates that fairness-aware AI in water governance cannot be evaluated through aggregate accuracy alone. It must be assessed through the distribution of errors across households and settlements that experience different forms of vulnerability.

The paper's central conceptual contribution is the infrastructural visibility gap. Water insecurity is not only a consequence of inadequate infrastructure; it is also shaped by what institutions measure, value, and act upon. Households may be formally visible as service recipients while remaining analytically invisible as bearers of time loss, income pressure, health risk, and coping burdens. AI can help reduce this gap only when household burden is incorporated into the data architecture and when model outputs are linked to institutional response protocols. Without such integration, AI may make water governance more efficient while preserving the invisibility of those most affected by unreliable, unsafe, or unaffordable access.

The broader implication is that the adoption of AI in urban water governance should not be treated as a technical transition alone. It is also a choice about what forms of evidence count in public decision-making. The challenge, therefore, is not simply whether utilities should adopt AI, but how the data architectures underpinning AI can be designed to reflect the lived realities of households most affected by water insecurity.

# Declarations

## Ethical Approval and accordance

The requirement of ethical approval for this secondary analysis was waived by the Research Ethics Review Committee / Internal Review Board of the Institute for Housing and Urban Development Studies (IHS), Erasmus University Rotterdam, in accordance with the Erasmus University Rotterdam research ethics review procedures and applicable institutional guidelines for research involving human participants.

No personally identifying household information is reported in this manuscript. All results are presented in aggregate form.

## Consent to participate

Informed consent was obtained from participants involved in the original household survey and fieldwork. Participation was voluntary, and participants were informed that results would be reported in aggregate form without personal identifiers.

## Consent to publish

Not applicable. This manuscript does not contain identifiable individual-level data, images, personal details, or case information requiring consent for publication.

## Data availability

The datasets generated during and/or analysed during the current study are not publicly available due to research ethics and participant confidentiality considerations but are available from the corresponding author on reasonable request.

## Conflict of interest

The author declares no conflict of interest.

## Funding

No external funding was received for this study.

## Clinical trial number

Not applicable.

## References


Anand, N. (2011). Pressure: The PoliTechnics of water supply in Mumbai. Cultural Anthropology, 26(4), 542–564.

Bakker, K., Kooy, M., Shofiani, N. E., & Martijn, E. J. (2008). Governance failure: Rethinking the institutional dimensions of urban water supply to poor households. World Development, 36(10), 1891–1915.

Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. California Law Review, 104, 671–732.

Chouldechova, A., Benavides-Prado, D., Fialko, O., & Vaithianathan, R. (2018). A case study of algorithm-assisted decision making in child maltreatment hotline screening decisions. Proceedings of the 1st Conference on Fairness, Accountability and Transparency, Proceedings of Machine Learning Research, 81, 134–148.

Eubanks, V. (2018). Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. St. Martin's Press.

Graham, S., & Marvin, S. (2001). Splintering Urbanism: Networked Infrastructures, Technological Mobilities and the Urban Condition. Routledge.

Howard, G., Ince, M. E., & Smith, M. D. (2003). Rapid Assessment of Drinking-Water Quality: A Handbook for Implementation. WEDC, Loughborough University.

Jepson, W. E., Wutich, A., Collins, S. M., Boateng, G. O., & Young, S. L. (2017). Progress in household water insecurity metrics: A cross-disciplinary approach. WIREs Water, 4(3), e1214.

Kitchin, R. (2014). The real-time city? Big data and smart urbanism. GeoJournal, 79(1), 1–14.

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366(6464), 447–453.

Palakodeti, A. (2021). S.I.T.E valuation of drinking water: A multi-disciplinary and integrated approach to provide safe and adequate drinking water to urban poor communities. Doctoral thesis, Erasmus University Rotterdam.

Ranganathan, M. (2014). Paying for pipes, claiming citizenship: Political agency and water reforms at the urban periphery. International Journal of Urban and Regional Research, 38(2), 590–608.

Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. Proceedings of the Conference on Fairness, Accountability, and Transparency, 59–68.

Stoler, J., Miller, J. D., Brewis, A., Freeman, M. C., Harris, L. M., Jepson, W., Pearson, A. L., Rosinger, A. Y., Shah, S. H., Staddon, C., Wutich, A., & Young, S. L. (2021). The Household Water Insecurity Experiences Scale: Comparison scores from 27 sites in 22 countries. Journal of Water, Sanitation and Hygiene for Development, 11(6), 1102–1110.

Taylor, L. (2017). What is data justice? The case for connecting digital rights and freedoms globally. Big Data & Society, 4(2), 1–14.

Young, S. L., Boateng, G. O., Jamaluddine, Z., Miller, J. D., Frongillo, E. A., Neilands, T. B., Collins, S. M., Wutich, A., Jepson, W. E., Stoler, J., & the HWISE Research Coordination Network. (2019). The Household Water Insecurity Experiences Scale: Development and validation of a household water insecurity measure for low-income and middle-income countries. BMJ Global Health, 4, e001750.

Zhu, M., Wang, J., Yang, X., Zhang, Y., Zhang, L., Ren, H., Wu, B., & Ye, L. (2022). A review of the application of machine learning in water quality evaluation. Eco-Environment & Health, 1(2), 107–116.