Eco-Environment & Health 5 (2026) 100209

Contents lists available at ScienceDirect

Eco-Environment & Health

journal homepage: www.journals.elsevier.com/eco-environment-and-health

Original Research Article

Integrated global assessment of *Escherichia coli* emissions from wastewater treatment plants

Wen Li[a,d,1], Qingbin Yuan[b,1], Xiaolong Wang[b], Wei Wang[b], Xinda Wu[b], Yanxu Zhang[c], Xiaoli Zhao[a], Yi Luo[b,*], Fengchang Wu[a,d,**]

[a] *State Key Laboratory of Environmental Criteria and Risk Assessment, Chinese Research Academy of Environmental Sciences, Beijing 100012, China*
[b] *State Key Laboratory of Water Pollution Control and Green Resource Recycling, School of the Environment, Nanjing University, Nanjing 210023, China*
[c] *School of Atmospheric Sciences, Nanjing University, Nanjing 210023, China*
[d] *State Key Laboratory of Pollution Control and Resource Reuse, College of Environmental Science and Engineering, Tongji University, Shanghai 200092, China*

ARTICLE INFO

Keywords:
Drinking water
Escherichia coli
Machine learning
Microbial discharge standards
Recreational water
Wastewater treatment plants

ABSTRACT

Wastewater treatment plants (WWTPs) are critical for mitigating anthropogenic contaminants, yet their effectiveness in controlling waterborne pathogens has not been assessed globally. This study presents the first global assessment of *Escherichia coli* (*E. coli*) emissions from 58,502 WWTPs across 188 countries and territories, utilizing a CatBoost machine learning model based on actual data from 230 facilities. Our results reveal pervasive *E. coli* contamination in global WWTP effluents, with a median concentration of 3.83 log Most Probable Number (MPN)/100 mL (range: 0–5.86 log MPN/100 mL), and 78.3% of WWTPs exceeding 3 log MPN/100 mL, particularly in Africa, South America, and Southeast Asia. The primary driver of this contamination is insufficient disinfection coverage, which is closely linked to the lack of enforceable microbial discharge standards rather than the development level alone. Furthermore, *E. coli* emissions from WWTPs compromise the microbial quality of downstream drinking and recreational waters, posing serious public health threats. To address this gap, we propose Human Development Index (HDI)-stratified *E. coli* discharge standards based on reverse risk modeling aligned with WHO health-based targets. Scenario analyses indicate that combining these risk-informed thresholds with targeted disinfection upgrades could reduce global non-compliance of WWTPs by 47% for drinking water and 58% for recreational water exposure, with the greatest benefits in regions with limited infrastructure. This study provides the first global, health-based framework for microbial water quality management in WWTPs, offering actionable guidance to reduce inequities in waterborne disease risk and improve global water safety.

1. Introduction

Wastewater treatment plants (WWTPs) play a vital role in mitigating anthropogenic pollution. Although they achieve substantial success in removing chemical contaminants, growing attention is being directed toward the potential health risks posed by microbial pathogens that may persist in treated effluents [1,2]. Discharged effluents typically enter rivers, lakes, and coastal waters that serve as essential sources of drinking water, recreational activities, and agricultural irrigation, thereby creating direct pathways for human exposure [3–6]. According to the World Health Organization (WHO), nearly half of Asia's rivers are heavily contaminated with fecal coliforms [7], and similar patterns have been observed in Sub-Saharan Africa [8,9].

Mounting evidence links inadequately treated wastewater to outbreaks of waterborne diseases, primarily through the ingestion of contaminated drinking water [3,10] and exposure during recreational water activities [4,5]. These risks are further exacerbated in settings where untreated or partially treated effluents are directly discharged into surface waters [11]. It is therefore increasingly important to quantify pathogen emissions from WWTPs and align them with human exposure routes for the design of effective, risk-based water safety interventions [8,9].

* Corresponding author.
** Corresponding author. State Key Laboratory of Environmental Criteria and Risk Assessment, Chinese Research Academy of Environmental Sciences, Beijing 100012, China.
E-mail addresses: luoy@nju.edu.cn (Y. Luo), wufengchang@vip.skleg.cn (F. Wu).
[1] The authors contributed equally to this work.

https://doi.org/10.1016/j.eehl.2025.100209
Received 22 May 2025; Received in revised form 31 August 2025; Accepted 15 December 2025
Available online 16 December 2025

Escherichia coli (*E. coli*) is widely recognized as a fecal indicator bacterium (FIB), owing to its predominant commensal origin in human and animal intestines and its ease of detection using standardized laboratory techniques [12,13]. Importantly, certain *E. coli* pathotypes, such as enterohemorrhagic, enterotoxigenic, and enteropathogenic strains, are known to cause gastrointestinal illness and therefore warrant particular attention in microbial water quality monitoring [14,15]. This dual role, serving as both an indicator of fecal contamination and an opportunistic pathogen, explains its widespread regulatory adoption in microbial water quality guidelines, including the WHO Guidelines for Drinking-Water Quality [16], the United States Environmental Protection Agency (U.S. EPA) Recreational Water Quality Criteria [17], and the European Union Bathing Waters Directive (BWD) [18].

Despite its global relevance, current *E. coli* monitoring data remain geographically fragmented, with limited coverage in low- and middle-income regions that are often the most vulnerable to waterborne disease risks [8,9,19]. Moreover, the lack of harmonized, large-scale microbial surveillance systems limits the ability to generalize localized findings to broader geographic contexts [3–5,10]. To address these limitations, emerging artificial intelligence techniques, particularly machine learning models, offer robust tools to integrate heterogeneous environmental data and predict microbial contamination across diverse spatial and socio-economic contexts, supporting evidence-based water quality management [20,21].

In this study, we present the first global assessment of *E. coli* contamination in WWTP effluents using a machine learning framework. The objectives were to: (i) quantify the magnitude and geographic distribution of *E. coli* discharges from WWTPs worldwide; (ii) evaluate their potential impacts on downstream drinking water sources and bathing water bodies; and (iii) establish health-based regional discharge standards and recommend feasible implementation measures. By providing a comprehensive and policy-relevant assessment of WWTP-derived *E. coli* emission risks, this work advances the global understanding of wastewater management challenges and offers a data-driven framework to strengthen water safety protections across varying infrastructural and socio-economic settings.

2. Materials and methods

2.1. Literature search and data collection

We conducted a comprehensive literature search on Web of Science to collect *E. coli* effluent concentration data from global WWTPs, which were used to train and validate the machine learning models. The search strategy employed the following query: TS = ("municipal sewage treatment plant" OR "wastewater treatment plant" OR "WWTPs") AND ("*Escherichia coli*" OR "*E. coli*") AND ("sewage" OR "sewages" OR "wastewater" OR "wastewaters" OR "inflow" OR "effluent"). To ensure objectivity in the selection process, two independent researchers conducted the screening and review of the articles. This initial search returned 3619 studies. After a preliminary screening of titles and abstracts, 892 studies were reviewed, of which 207 were retained after further scrutiny. Detailed inclusion criteria were applied to select reports that provided: (i) quantitative measurements of thermotolerant coliforms or *E. coli* in both WWTP influents and effluents, used as indicators of fecal contamination; and (ii) essential WWTP information, with details such as their geographical locations, populations served, daily wastewater discharges, treatment levels, and disinfection status. This process resulted in the compilation of 230 distinct data points from 70 individual studies, spanning 37 countries across six WHO regions, with data covering the period from 2000 to 2023 (Fig. S1, Table S1). Data were extracted from text, tables, and figures, with graphical data digitized using Plot Digitizer software (http://plotdigitizer.sourceforge.net/). To minimize potential data representation bias, a data representation analysis was carried out to ensure that the training dataset was geographically diverse and adequately represented countries across various Human Development Index (HDI) categories (Fig. S2, Text S1).

The model incorporates eight key input features, identified based on expert judgment, that significantly influence *E. coli* concentrations in WWTP effluents. These features include: (i) climate factors, consisting of annual average temperature and total precipitation, which influence WWTP operations and *E. coli* concentrations in effluents [12,22]; (ii) population loads, represented by human daily fecal production and the population served by WWTPs, which affect influent characteristics and serve as a primary source of *E. coli* in WWTPs [23,24]; and (iii) wastewater treatment conditions, comprising daily wastewater discharges, disinfection status, and treatment levels (split into two binary variables indicating the presence of secondary and advanced treatment, respectively), which are critical for assessing *E. coli* removal [25,26]. Climate data were sourced from the ERA5 monthly-averaged dataset [27], publicly available at https://cds.climate.copernicus.eu, and were processed in ArcGIS 10.8 using WWTP coordinates and dates from the training database. Human daily fecal production was calculated using annual data [24]. Spearman's rank correlation analysis revealed a significant relationship between the input features and the predicted *E. coli* concentration (Fig. S4).

2.2. Construction of machine learning models

Nine commonly used regression algorithms were implemented using the scikit-learn library in Python to predict *E. coli* concentrations (**Text S3**). These algorithms include Support Vector Regression (SVR), K-Nearest Neighbors (KNN), Decision Tree (DT), Random Forest (RF), CatBoost, XGBoost, LightGBM (LGBM), Gradient Boosting Decision Tree (GBDT), and Artificial Neural Network (ANN) (Fig. S1, Text S3). Data preprocessing involved standardization and necessary unit-type conversions to address scale variations across features and ensure model compatibility (Text S2).

To minimize bias in small data regimes and enhance model robustness, nested cross-validation was employed, with hyperparameter optimization using the Optuna framework (Texts S4 and S5). The inner loop utilized 3-fold cross-validation for tuning, while the outer loop used 5-fold cross-validation. The optimal hyperparameters, selected based on the lowest root-mean-square error (RMSE) from the inner loop, were applied to train the model in the outer fold and evaluate its performance on the test set (Texts S4 and S5). This structure effectively prevented data leakage, mitigated overfitting, and enhanced generalization, particularly for small datasets [28]. Additionally, model reliability was further validated through multiple randomized data splits with ten random seeds [29].

2.3. Evaluation of model performance

Model performance was assessed using the coefficient of determination (R^2), RMSE, and mean absolute error (MAE), with lower RMSE and MAE values and higher R^2 values indicating stronger predictive accuracy (Text S6). To improve interpretability, Shapley additive explanations (SHAP) were applied to quantify the contribution of individual features and identify key predictors (Text S7). The overall analysis framework is illustrated in Fig. S1. Among the nine algorithms evaluated, the optimized CatBoost model achieved the highest predictive performance, with an RMSE of 0.80, an MAE of 0.65, and an R^2 of 0.80 on the test set (Table S6). Nested cross-validation further confirmed its robustness, yielding an RMSE of 0.77 ± 0.17, an MAE of 0.63 ± 0.16, and an R^2 of 0.76 ± 0.10 (Table S6). Validation across ten random seed splits produced similarly consistent results (RMSE: 0.77–0.80; MAE: 0.62–0.65; R^2: 0.74–0.76), underscoring the model's stability and generalizability (Fig. S5).

CatBoost's superior performance stems from its ability to natively process categorical variables, such as treatment level and disinfection status, without requiring manual encoding, thereby preserving feature structure in heterogeneous environmental datasets [30]. In contrast to algorithms like RF and XGBoost, which rely on extensive preprocessing, CatBoost combines efficient categorical handling with built-in regularization and symmetric tree construction, resulting in improved performance on small and imbalanced datasets [31,32]. This methodological advantage is reflected in stronger predictive outcomes compared to previous models of microbial contamination. For example, an SVR model trained on 1503 samples from 79 WWTPs achieved an R^2 of 0.63 [33], while an ANN model using 32–74 samples from 3 WWTPs reported R^2 values between 0.60 and 0.79 [20]. By comparison, CatBoost demonstrates clear potential for large-scale predictive modeling, particularly in data-limited or feature-diverse scenarios [20,34].

2.4. Model application

After developing the final predictive model, CatBoost, which demonstrates stable performance, we applied it to predict effluent *E. coli* concentrations for 58,502 WWTPs globally. The basic information, locations, and operational parameters of these WWTPs were sourced from the HydroWASTE database [35]. However, this dataset lacked disinfection status data. To address this, we supplemented the dataset with disinfection details from the U.S. EPA and European Environment Agency (EEA) databases, covering 65.8% of the WWTPs (Fig. S8). For the remaining missing data, a classification model trained on 230 actual WWTP records was used to predict disinfection status (Text S8). To maintain consistency with the model development, all data from the HydroWASTE database were standardized and aligned based on WWTP coordinates and country names. The final CatBoost model, implemented in Python, was used to generate the global spatial distribution of *E. coli* concentrations across the 58,502 WWTPs in 188 countries and territories. Daily *E. coli* discharge loads by country were estimated to assess their contribution to the global burden on aquatic environments (Fig. S12, Text S9).

2.5. Reverse modeling to derive HDI-stratified E. coli discharge standards

To derive science-based *E. coli* discharge standards, we employed a reverse modeling approach grounded in international health guidelines and calibrated for regional implementation capacity. For drinking water protection, we followed the WHO recommendation of at least a 2.6-log reduction in pathogen concentrations through conventional treatment processes in waterworks [16]. Assuming non-detectable E. coli levels in finished water and incorporating river dilution coefficients, we back-calculated acceptable WWTP effluent thresholds. For recreational water protection, we adopted conservative national benchmarks, 2.70 log Most Probable Number (MPN)/100 mL (i.e., 500 MPN/100 mL from Morocco) for low- and medium-HDI countries and 2.10 log MPN/100 mL (i.e., 126 MPN/100 mL from U.S. EPA) for high- and very high-HDI countries, and similarly applied in-river dilution to derive effluent-based thresholds (Table S8). To reflect regional differences in regulatory and infrastructural capacity, all thresholds were stratified by HDI level (Fig. S13). Initial compliance targets were set at 95%, then relaxed to 50% for low- and medium-HDI regions to align with the United Nations Sustainable Development Goal (SDG) 6.3 of halving untreated wastewater discharge by 2030 [36]. The economic feasibility of the proposed standards was also evaluated by estimating the annualized cost of installing chlorination systems, expressed as a percentage of national GDP (Fig. S15, Text S10).

2.6. Statistical analysis

To visualize the data distribution of *E. coli* concentrations in global WWTP effluents, national daily *E. coli* loads, disinfection practices, and related factors, Geographic Information System (GIS) analyses were performed using ArcGIS 10.8. The GEMStat Water Quality database (https://gemstat.org/data-gemstat/data-portal/) and bathing water quality data from European Union member states (https://www.eea.europa.eu/en/topics/in-depth/bathing-water/state-of-bathing-water) were utilized to assess the potential impact of *E. coli* emissions from WWTPs on aquatic environments. GIS was used to identify WWTPs located within 30 km upstream of European Union bathing sites and GEMStat water quality monitoring locations, a conservative distance that represents the area where WWTPs may contribute to microbial contamination in the surrounding watershed [37].

For statistical analysis, probability density and cumulative distribution, Sankey diagrams, one-way analysis of variance (ANOVA), linear regression, and correlation analyses were performed using Origin 10.5 (OriginLab Corp., Northampton, MA, U.S.) and its official plugins. Detailed descriptions of these analyses are provided in the figure legends for each experiment. Partial least squares path modeling (PLS-PM) was performed using the R plspm package.

3. Results

3.1. Global emissions of E. coli from WWTPs

This study presents a comprehensive predictive dataset of *E. coli* concentrations in the effluents of 58,502 WWTPs across 188 countries and territories. Predicted values range from 0 to 5.86 log MPN/100 mL, with a global median of 3.83 log MPN/100 mL, closely aligning with the distribution of measured values from global observations (0–6.60 log MPN/100 mL) (Figs. 1 and S3a, Table S1). Notably, 78.3% of WWTPs exceed 3 log MPN/100 mL (i.e., 1000 MPN/100 mL), a commonly used microbial threshold (Fig. 1, Table S7).

Predicted *E. coli* concentrations display pronounced regional heterogeneity, with elevated levels observed in many low-HDI regions, including parts of Africa, South America, and Southeast Asia (Figs. 1 and S10). The African region (AFRO) reports the highest *E. coli* concentrations across all six WHO regions, averaging (4.58 $\pm$ 0.66) log MPN/100 mL (Fig. S10b). Hotspot areas include Zambia [(5.59 $\pm$ 0.17) log MPN/100 mL], Rwanda [(5.22 $\pm$ 0.10) log MPN/100 mL], and several Southeast Asian countries such as Bangladesh [(4.92 $\pm$ 0.13) log MPN/100 mL] and Nepal [(4.88 $\pm$ 0.61) log MPN/100 mL] (Figs. 1 and S10a). Medium-HDI countries in the Americas, such as Haiti, also report elevated concentrations [(4.83 $\pm$ 0.24) log MPN/100 mL]. Many high-HDI countries report relatively low concentrations, for example, Bahrain [(0.62 $\pm$ 0.02) log MPN/100 mL] and Singapore [(2.37 $\pm$ 1.51) log MPN/100 mL] (Figs. 1 and S10a), likely reflecting the benefits of well-developed wastewater infrastructure and consistent enforcement of water safety regulations [38].

However, notable exceptions persist. The European Region average [(3.77 $\pm$ 0.85) log MPN/100 mL] is the highest among non-African WHO regions, and Australia reports a high average of (4.33 $\pm$ 0.57) log MPN/100 mL despite its advanced infrastructure (Figs. 1 and S10). Moreover, the average concentration of high-HDI countries is significantly ($p < 0.05$) higher than that of medium-HDI countries (Figs. 1 and S11b). In contrast, a few low- to medium-HDI countries show comparatively low effluent *E. coli*, for example, Sudan [(1.51 $\pm$ 0.04) log MPN/100 mL] and Morocco [(3.68 $\pm$ 1.15) log MPN/100 mL] (Figs. 1 and S10a). Taken together, these patterns suggest that HDI alone does not fully account for microbial outcomes in wastewater treatment and underscore a partial decoupling between development level and treatment performance [9,39]. Previous studies have shown that effective disinfection strategies, such as chlorination, ultraviolet, or ozone treatment, can reduce *E. coli* concentrations by 1–6 log units [40–42]. Therefore, while high HDI may indicate greater technical capacity, consistent microbial control ultimately depends on effective regulatory enforcement and operational execution.

3.2. Insufficient disinfection as the primary driver of E. coli release

The critical role of disinfection in reducing *E. coli* contamination in WWTP effluents is strongly supported by both machine learning and

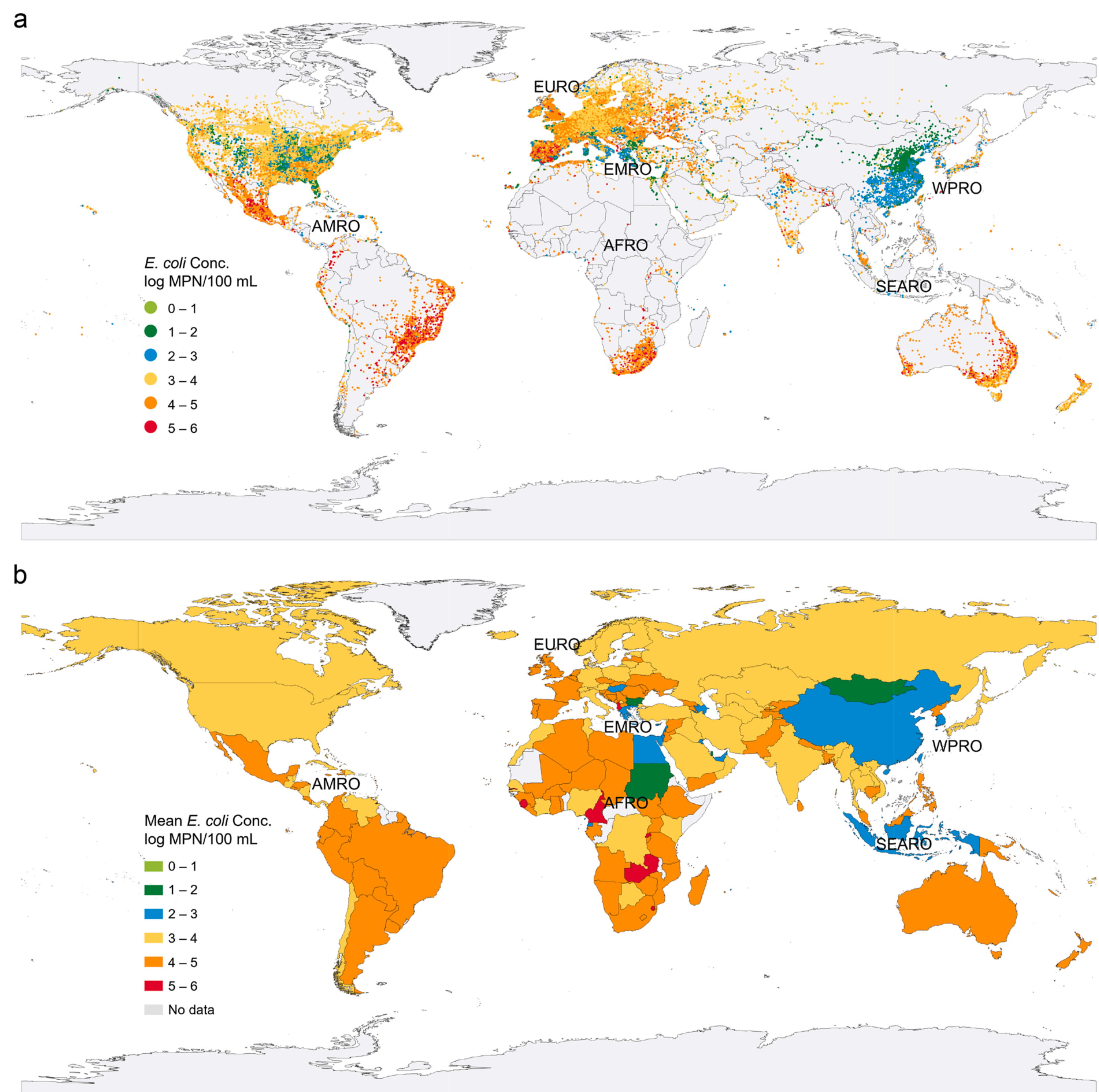

Fig. 1. Global distribution of *E. coli* contamination in WWTP effluents. (a) Global map showing *E. coli* concentrations in the effluents of 58,502 WWTPs. Each point represents a facility, with concentrations indicated by a color gradient from green (low: 0 log MPN/100 mL) to red (high: 5.86 log MPN/100 mL). (b) Average *E. coli* concentrations in WWTP effluents by country. AFRO, African Region; EMRO, Eastern Mediterranean Region; EURO, European Region; AMRO, Region of the Americas; SEARO, South-East Asia Region; WPRO, Western Pacific Region.

statistical analyses. Feature importance analysis revealed that disinfection status exerts a greater influence on effluent *E. coli* concentrations than other factors such as climate or population density (Fig. 2a). This result was further validated by PLS-PM, which demonstrated a significant ($p < 0.05$) reduction in *E. coli* levels associated with disinfection in both the training dataset (n = 230) and the global set of predicted WWTPs (n = 58,502) (Fig. 2b–d). Consistent with these findings, countries with higher disinfection coverage generally exhibit lower effluent *E. coli* concentrations (Fig. S9e), underscoring inadequate disinfection as a major driver of elevated microbial pollution in many regions.

National variation in disinfection coverage is shaped by differences in both economic capacity and regulatory infrastructure (Figs. 2e and S9d). In lower-income regions, financial and infrastructural constraints are major barriers; for example, only 9.3% of WWTPs in Africa and 7.2% in South America are equipped with disinfection systems (Fig. S9). While development level influences infrastructure availability, enforceable microbial discharge standards more directly drive disinfection implementation, as shown by significantly higher coverage in countries with such regulations (Fig. 2e). Currently, only 25 countries, including China, India, Kenya, Belize, and several nations in Europe and Latin America, have established microbial discharge standards (Fig. 2e,

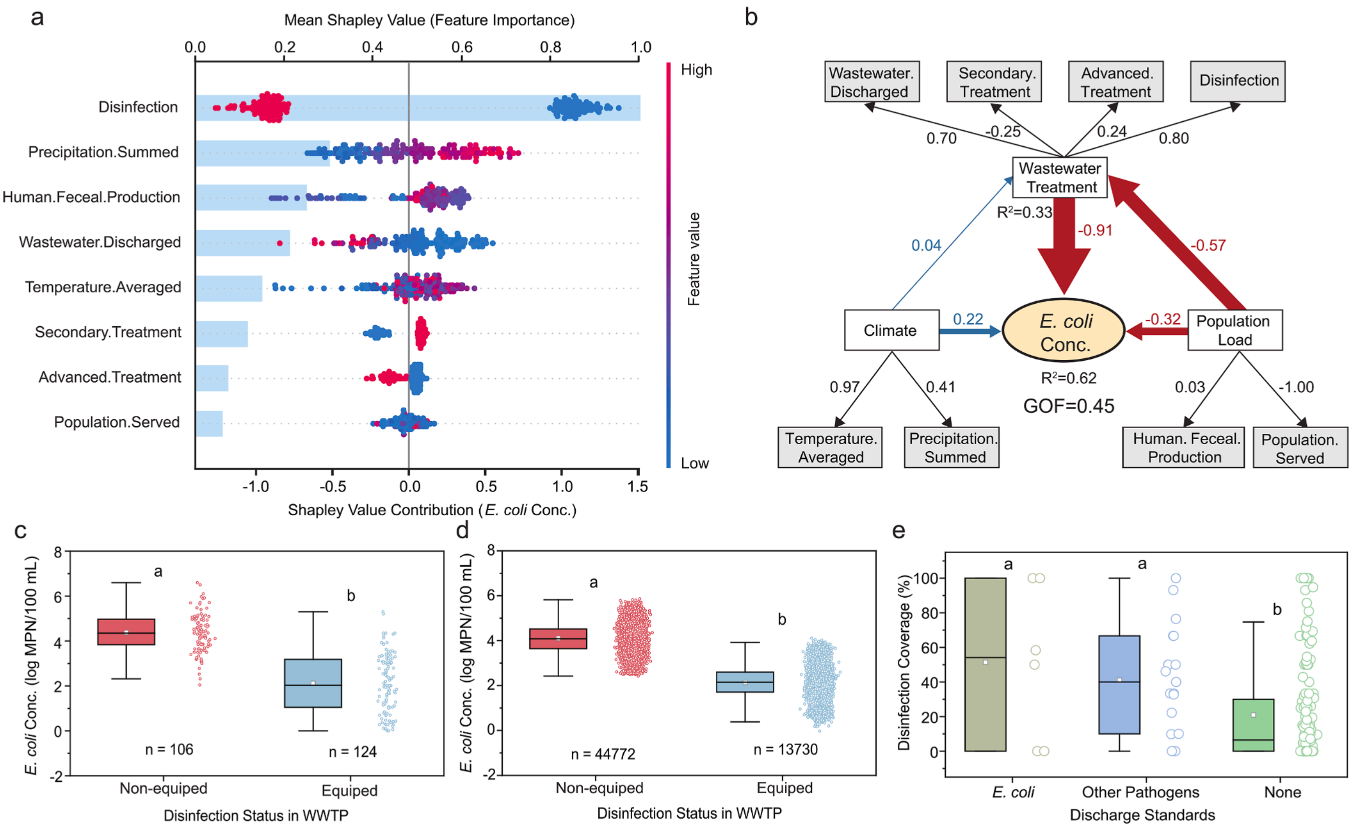

Fig. 2. Influence of disinfection on *E. coli* emissions from global WWTPs. (a) Shapley additive explanations (SHAP) for CatBoost regression model features. Each row represents a feature, and each point denotes the Shapley value for an individual sample. Negative Shapley values indicate a negative impact on *E. coli* concentration, and positive values indicate a positive impact. The color gradient ranges from blue (lower feature values) to red (higher feature values). The bar plot displays the average absolute Shapley value for each feature, highlighting disinfection as the most significant factor influencing *E. coli* concentrations in WWTP effluents. (b) Partial least squares path modeling (PLS-PM) showing the direct and indirect effects of climate parameters, population load, and wastewater treatment conditions on *E. coli* concentrations. Solid lines represent significant paths, blue lines show positive correlations, and red lines indicate negative correlations. Line thickness reflects correlation strength. Numbers on the lines indicate path coefficients, and numbers on the black lines represent the loadings of indicators on observed variables. Disinfection practices emerge as the most influential factor. Comparison of effluent *E. coli* concentrations by disinfection status, based on 230 observed (c) and 58,502 predicted (d) global WWTPs data. (e) Relationship between regulatory discharge standards and national disinfection coverage. Different lowercase letters denote statistically significant differences ($p < 0.05$).

Table S7). The benefits of regulation are evident, for example, in China, the Discharge Standard of Pollutants for Municipal Wastewater Treatment Plants sets a fecal coliform limit of 3 log MPN/100 mL, resulting in 93.1% disinfection coverage and low effluent *E. coli* levels [(2.19 ± 0.57) log MPN/100 mL], ranking 182 out of 188 countries and territories (Figs. S9b and S10a, Table S7). Similarly, India's identical standard has supported 66.7% disinfection coverage and moderately reduced *E. coli* levels [(3.76 ± 0.82) log MPN/100 mL], slightly below the global median of 3.83 log MPN/100 mL (Figs. S9b and S10a, Table S7).

Conversely, even high-HDI countries with extensive WWTP infrastructure may report low disinfection coverage in the absence of enforceable microbial discharge standards. The European Union and the U.S. together account for 86.1% of global WWTPs (Fig. 1a), yet disinfection coverage remains limited, only 19.6% (EEA) and 22.7% (U.S. EPA), respectively, primarily due to the lack of microbial standards (Fig. S8, Table S7). As a result, *E. coli* levels in treated effluents remain elevated, averaging (3.77 ± 0.85) log MPN/100 mL in the European region and (3.10 ± 1.09) log MPN/100 mL in the U.S. (Figs. 1 and S10). These findings suggest that, for effective microbial control in wastewater treatment, regulatory frameworks may be more critical than infrastructure alone, and that this relationship cannot be fully explained by national income levels.

3.3. Global *E. coli* emissions pose alarming health concerns

Insufficient disinfection in WWTPs contributes substantially to the global *E. coli* burden in receiving waters, with daily discharge loads ranging from $10^{8.0}$ to $10^{14.7}$ CFU *E. coli* per country (Fig. S12, Text S9). These elevated levels and discharge loads pose direct threats to public health, particularly in regions that rely on pathogen-contaminated water sources for drinking and recreational purposes [3–5].

The threat to drinking water safety is particularly pronounced given that approximately two-thirds of the global supply is sourced from surface water, which is often the primary recipient of WWTP effluents [19]. *E. coli* concentrations in WWTP discharges show a significant positive correlation ($r = 0.32$, $p < 0.01$) with the proportion of freshwater drawn from surface water (Fig. 3a), indicating elevated exposure risks in countries with high surface water dependence. For instance, Brazil, sourcing 88.4% of its freshwater from surface water, reports elevated *E. coli* levels in WWTP effluents [(4.64 ± 0.49) log MPN/100 mL], well above the global average of 3.64 log MPN/100 mL (Fig. 3a). Even in countries with lower surface water reliance, such as Denmark (21.8%), effluent concentrations remain concerning (3.86 log MPN/100 mL) (Fig. 3a). Furthermore, *E. coli* levels in WWTP effluents correlate significantly ($r = 0.31$, $p < 0.001$) with those in downstream water bodies within 30 km (Fig. 3b), with contamination extending to

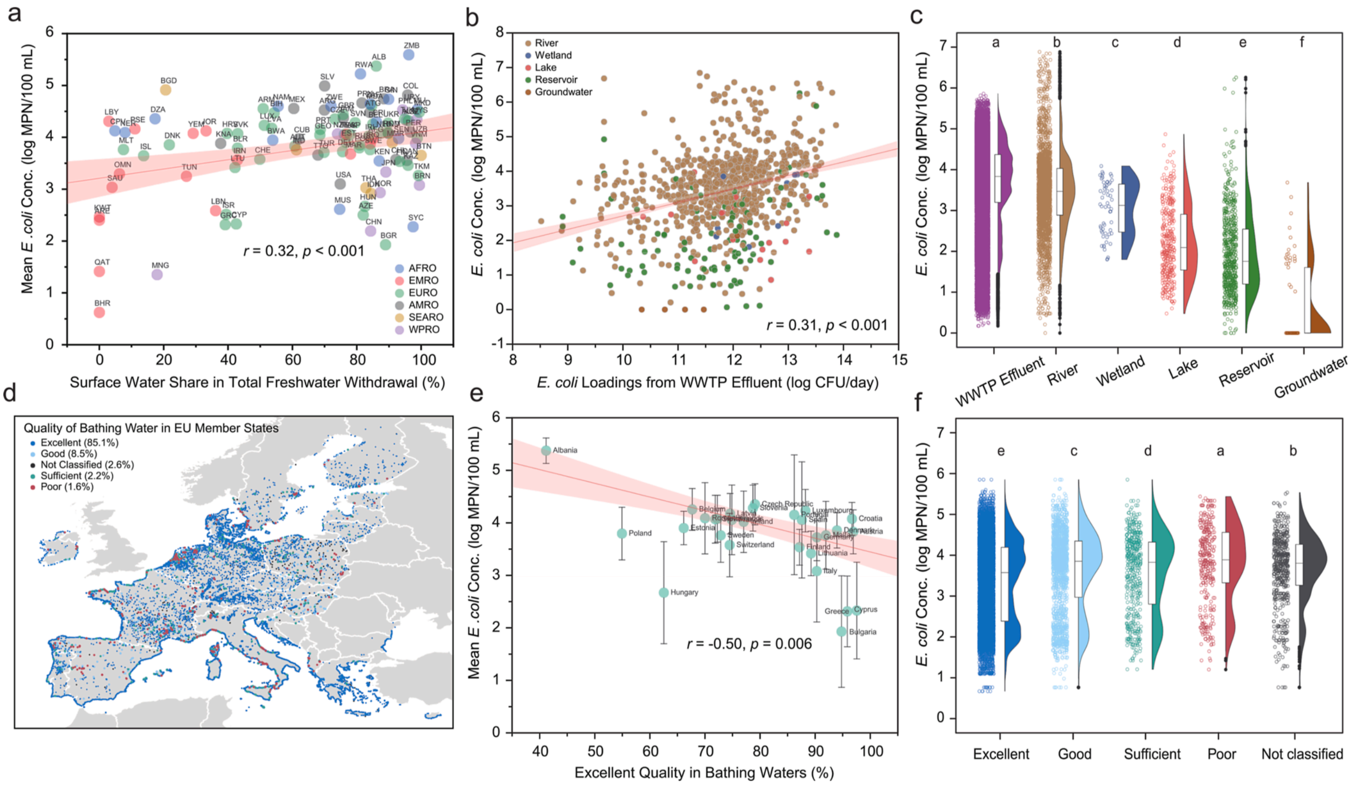

Fig. 3. Impact of *E. coli* emission from WWTPs on surface water quality. (a) Linear regression between the proportion of surface freshwater in total freshwater withdrawal(https://www.fao.org/faolex/zh/) and average *E. coli* concentrations in WWTP effluents across countries. The significant positive correlation ($r = 0.32$, $p < 0.001$) indicates that countries relying more on surface water face higher risks of *E. coli* contamination. The red-shaded areas represent 95% confidence intervals (CI). The three-letter ISO codes represent countries and regions (https://www.iso.org). (b) Linear regression between *E. coli* in natural waters and daily loads from upstream WWTPs (within 30 km, a conservative range for potential microbial contamination). The positive correlation ($r = 0.31$, $p < 0.001$) suggests that WWTP discharges influence *E. coli* levels in river systems. Red-shaded areas represent 95% CI. (c) Box plot comparing predicted *E. coli* concentrations in WWTP effluents with observed values in natural water bodies from the global GEMStat water quality database (https://gemstat.org/data-gemstat/data-portal). Different lowercase letters indicate significant differences between groups ($p < 0.05$). (d) Bathing water quality in European Union member states (https://www.eea.europa.eu). (e) Linear regression between the proportion of excellent bathing water quality and average *E. coli* concentrations in European Union bathing sites. The significant negative correlation ($r = -0.50$, $p = 0.006$) underscores the impact of WWTP emissions on bathing water quality. Red-shaded areas represent 95% CI. (f) Comparison of microbial water quality grades and *E. coli* concentrations in WWTP effluents within 30 km of bathing sites in the European Union. Different lowercase letters indicate significant differences between groups ($p < 0.05$).

drinking water reservoirs (1.95 log MPN/100 mL) and even groundwater (0.57 log MPN/100 mL) (Fig. 3c). These findings underscore the susceptibility of both surface and groundwater sources to fecal contamination, especially in low- and middle-income countries with limited wastewater treatment infrastructure [8,22,43]. However, such risks extend beyond developing regions, with high-income countries likewise experiencing repeated incidents of drinking water contamination [44]. For example, groundwater contamination from sewage and surface runoff caused multiple outbreaks in the U.S. , the United Kingdom, and Finland [45–47]. Between 1988 and 2002, Finland alone reported an estimated 15,000 waterborne illness cases [45,46]. From 2000 to 2014, 11 groundwater-related disease outbreaks were reported in 15 developed countries, including the U.S. and the United Kingdom, resulting in over 10,000 cases of gastrointestinal illness [47].

Recreational waters are similarly affected. Globally, over 120 million cases of gastrointestinal diseases and 50 million cases of acute respiratory diseases are attributed to swimming in contaminated waters [48]. Our results show that *E. coli* concentrations in effluent directly influence water quality in nearby recreational areas, with high effluent concentrations in regions like Europe correlating with poor recreational water quality (Fig. 3d–f). A significant negative correlation ($r = -0.50$, $p = 0.006$) between *E. coli* concentrations in WWTP effluents and the proportion of "excellent" bathing waters further highlights the role of effluent contamination in deteriorating recreational water quality (Fig. 3d,e). Poorer bathing water quality is associated with higher effluent concentrations (3.79 log MPN/100 mL) (Fig. 3f). These findings underscore the ongoing risk of contamination in recreational waters due to *E. coli* emissions from global WWTPs.

More broadly, *E. coli* contamination is widespread across global freshwater systems, with average concentrations in rivers (3.46 log MPN/100 mL), wetlands (3.05 log MPN/100 mL), and lakes (2.23 log MPN/100 mL), all significantly exceeding the U.S. EPA's recreational water quality threshold of 2.10 log MPN/100 mL (Fig. 3c). These elevated concentrations highlight the serious public health risks associated with inadequate wastewater treatment. For example, during 2003–2004, 62 recreational water-associated outbreaks were reported across 26 states and Guam in the U.S., resulting in 2698 illnesses, 58 hospitalizations, and 1 death [49]. Similarly, the Centers for Disease Control and Prevention documented 119 outbreaks linked to untreated recreational waters across 31 states between 2009 and 2019, leading to more than 5240 cases, predominantly during summer months, with *E. coli* implicated in 19 outbreaks involving 240 cases [50]. These findings underscore the urgent need for improved wastewater treatment and for implementing more stringent regulatory frameworks to mitigate microbial contamination and safeguard public health.

3.4. Science-based discharge standards for risk mitigation

As highlighted earlier, widespread *E. coli* contamination in WWTP effluents is largely attributable to insufficient disinfection coverage and the absence of enforceable microbial discharge standards. To address this gap, we propose a set of science-based *E. coli* discharge standards stratified by HDI level, incorporating both exposure scenarios and regional implementation capacity. This tiered framework aims to safeguard drinking and recreational water sources in a way that is both regionally adaptable and economically feasible (Fig. 4).

For drinking water source protection, we adopted a health-based target approach following the WHO Guidelines for Drinking-water Quality, which recommend a minimum 2.6-log reduction of pathogenic bacteria across the drinking water treatment process [16]. Using reverse modeling, we reconstructed source water *E. coli* concentrations and incorporated in-river dilution effects to estimate WWTPs effluent thresholds necessary to achieve non-detectable levels in finished drinking water [16]. To reflect regional variability in baseline effluent quality and compliance capacity, we derived HDI-stratified discharge standards of 4.49, 4.84, 2.84, and 3.67 log MPN/100 mL for low-, medium-, high-, and very high-HDI countries, respectively (Fig. 4a–d). These thresholds are projected to achieve approximately 95% compliance in high- and very high-HDI countries and around 50% compliance in low- and medium-HDI countries, in line with SDG 6.3, which targets a 50% reduction in untreated wastewater globally by 2030 [36]. A similar framework was applied to recreational water protection (Fig. 4e–h). Conservative microbial quality criteria were selected based on the most stringent national standards available: Morocco's inland water standard (2.70 log MPN/100 mL) for low- and medium-HDI countries, and the U.S. EPA's freshwater standard (2.10 log MPN/100 mL) for high- and very high-HDI countries (Table S8). After accounting for typical in-river dilution, we derived corresponding effluent standards of 4.59, 4.94, 2.34, and 3.17 log MPN/100 mL for low-, medium-, high-, and very high-HDI countries, respectively (Fig. 4e–h). These values align microbial health protection goals with region-specific exposure conditions and implementation capacities.

A global assessment of *E. coli* levels in WWTP effluents against the proposed discharge standards reveals widespread non-compliance (Fig. 5a,b). Over half of WWTPs currently exceed discharge limits for drinking water and recreational water protection, with 57% and 76% of facilities failing to meet the respective standards, respectively (Figs. 5a,b and S14a). Notable hotspots were observed in the AFRO, South America, and parts of Europe and Australia, where over 80% of WWTPs exceeded the proposed standards, particularly for recreational water protection (Fig. 5a,b). To mitigate this risk, we recommend coupling the implementation of discharge standards with targeted facility upgrades, especially the installation of disinfection systems in facilities that lack them and currently exceed the standards. With these interventions, global disinfection coverage is projected to increase by 56% for drinking water protection and by 73% for recreational water protection, with pronounced improvements in AFRO and South America. In some countries, disinfection coverage is expected to increase by more than 80% (Fig. 5c,d). These upgrades are projected to substantially reduce non-compliance, with only 10% and 18% of WWTPs remaining above the discharge standards for drinking and recreational water protection, respectively (Fig. S14a). Overall, the global exceedance rate is expected to decrease by 47% and 58% under drinking and recreational water protection, respectively (Fig. 5e,f). Across regions such as AFRO and EURO, most countries are projected to see declines of more than 30% in the share of WWTPs exceeding the standards (Fig. 5e,f). Furthermore, the proportion of WWTPs with *E. coli* concentrations exceeding 5 log MPN/100 mL declined from 42% to 37% for drinking water protection and from 42% to 20% for recreational water protection (Fig. S14b). These findings highlight the feasibility of HDI-stratified discharge standards combined with disinfection interventions as a flexible strategy for microbial risk reduction. By aligning health protection goals with the capacity for implementation, this approach offers a realistic and globally scalable pathway toward safer wastewater management and more equitable water quality protections.

4. Discussion

We addressed key challenges in model development and result reliability, particularly those arising from limited data availability and missing variables. Although the final training dataset included only 230 WWTPs, these were drawn from 37 countries that collectively account for 82.3% of the global WWTPs distribution across all six WHO regions, ensuring broad geographic and HDI-level representativeness (Fig. S2). To address gaps in disinfection records for 34.2% of the 58,502 WWTPs worldwide, we employed a CatBoost classification model that achieved strong predictive performance (accuracy: 0.87; precision: 0.88; AUC: 0.94) and external validity (19/21 correct predictions) (Figs. S6-S7, Table S4, Text S8). This imputation approach follows standard procedures widely used in machine learning research [51]. Coupled with

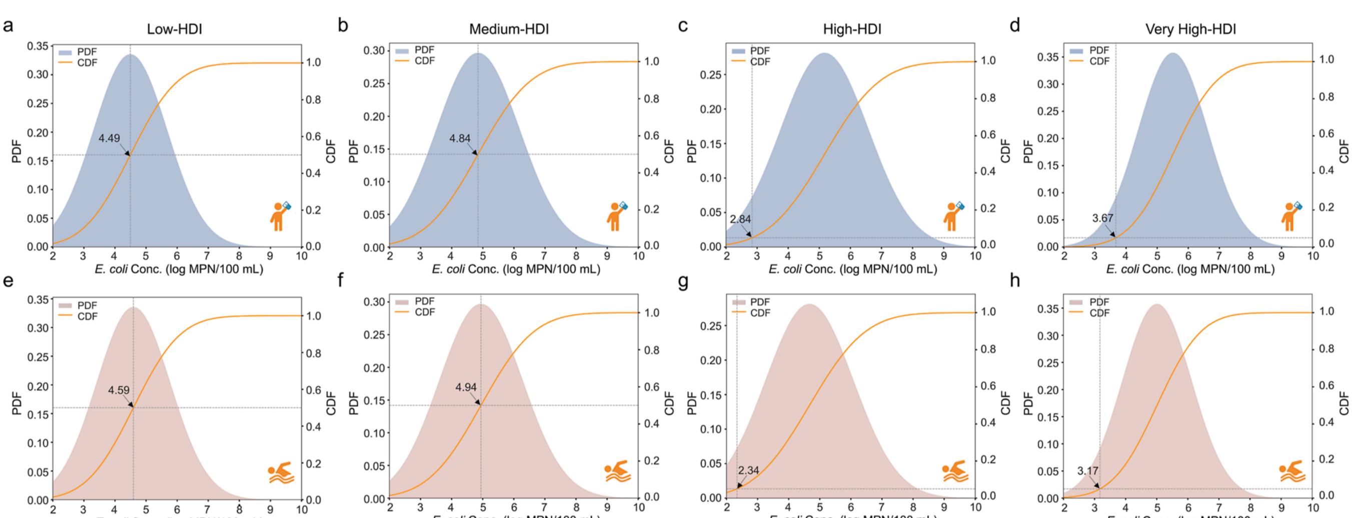

Fig. 4. Derivation of HDI-stratified *E. coli* discharge standards using reverse risk modeling. (a–d) Probability density and cumulative distribution curves of predicted WWTP effluent *E. coli* concentrations inferred from the WHO health-based target for drinking water safety (non-detectable in finished water). (e–h) Corresponding curves based on conservative recreational water thresholds: 2.30 log MPN/100 mL (Morocco) for low- and medium-HDI countries, and 2.10 log MPN/100 mL (U.S. EPA) for high- and very high-HDI countries. Arrows indicate proposed discharge standards achieving 95% WWTPs compliance in high- and very high-HDI countries and 50% in low- and medium-HDI countries, aligned with SDG 6.3, which aims to halve the proportion of untreated wastewater discharges by 2030.

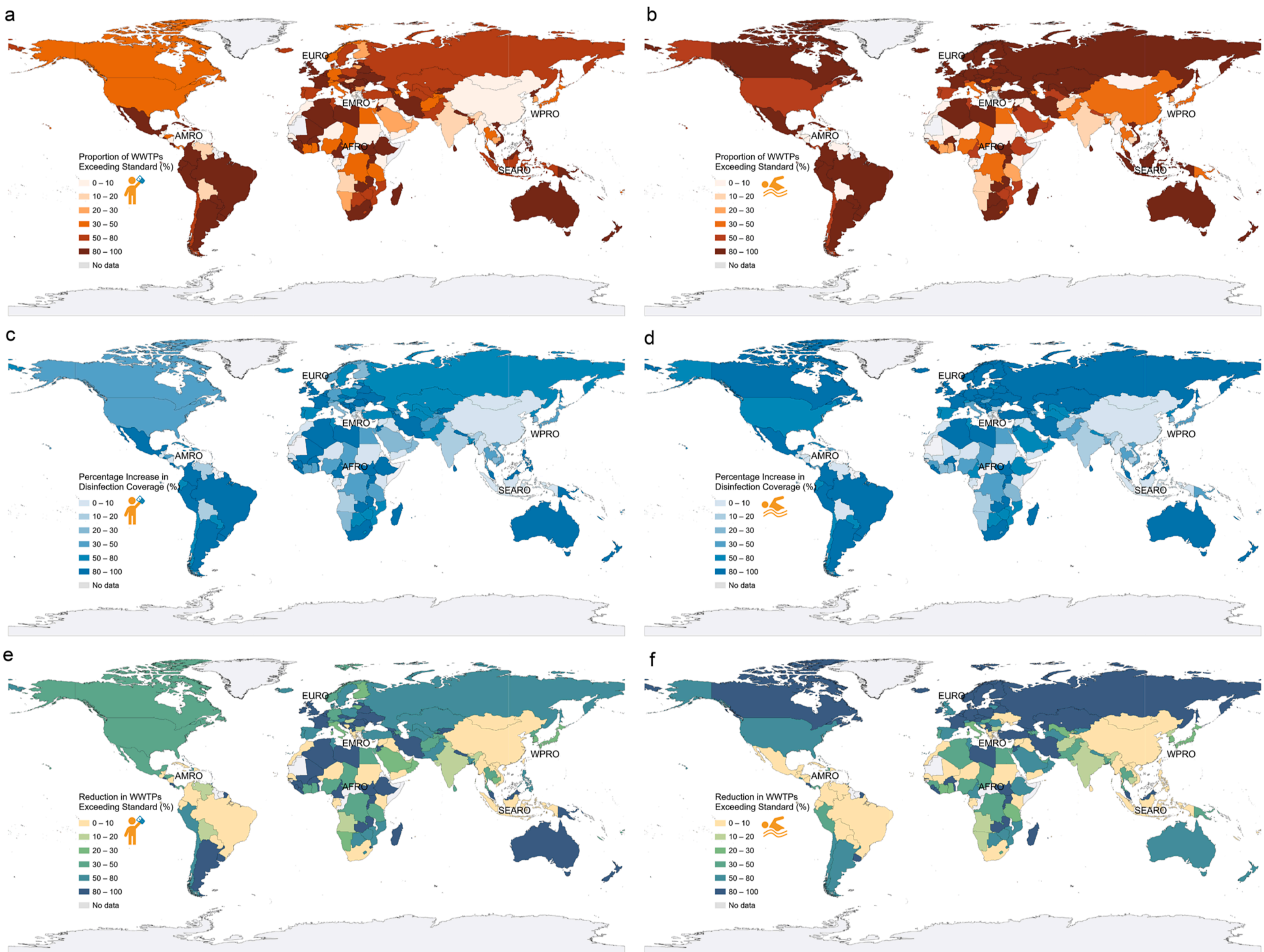

Fig. 5. Performance of HDI-stratified *E. coli* discharge standards and optimized disinfection strategies across global WWTPs. Proportion of WWTPs exceeding proposed discharge standards for drinking (a) and recreational water protection (b). Projected increases in the proportion of disinfection coverage following disinfection optimization, based on proposed discharge standards for drinking (c) and recreational water protection (d). Projected reductions in the proportion of non-compliant WWTPs after disinfection upgrades, based on proposed discharge standards for drinking (e) and recreational water protection (f). All analyses are presented at the country level.

nested cross-validation and repeated random splits (Texts S4 and S5), particularly suited to small datasets, these measures enhanced model robustness and ensured reliable extrapolation globally [28].

Building on these robust modeling efforts, our results reveal widespread and persistent *E. coli* emissions from WWTPs globally, primarily driven by insufficient disinfection and the absence of enforceable microbial discharge standards (Figs. 1 and 2). This regulatory gap leaves downstream drinking and recreational waters vulnerable to fecal contamination, thereby increasing public health risks (Fig. 3). To address this, we proposed a tiered discharge framework stratified by HDI level, grounded in WHO health-based targets and adjusted for exposure scenarios, receiving water dilution, and infrastructure capacities (Figs. 4 and S13). A 95% compliance target was set for high- and very high-HDI countries, while a 50% target was adopted for low- and medium-HDI countries to provide a realistic minimum threshold that aligns with SDG 6.3 of halving untreated wastewater discharge by 2030 [36].

To facilitate practical implementation, particularly in resource-limited regions, we recommend a risk-based approach that prioritizes disinfection upgrades at WWTPs that lack disinfection and exceed discharge standards. Model simulations show that such targeted interventions could reduce the number of non-compliant WWTPs globally by 47% and 58% for drinking and recreational water protection, respectively (Fig. 5e,f). Cost-effectiveness analyses further show that such interventions would require a modest investment, only 0.0113% and 0.0131% of national GDP on average for drinking water and recreational water protection, respectively (Fig. S15), well below the typical 0.5% GDP investment for the water sector in developing countries [52]. Given heterogeneous fiscal capacities, disinfection technology choices should be commensurate with local economic conditions. Although advanced technologies such as ultraviolet and ozone systems offer higher *E. coli* removal efficiencies (Fig. S11a), chlorination remains the most practical and cost-effective option, especially in low-income regions [53,54]. Additionally, targeted strategies, such as seasonal disinfection during high-exposure periods, already adopted in Europe and Latin America [55–57], could further reduce costs without sacrificing public health protection. By focusing on the high-risk facilities and leveraging locally appropriate technologies, this targeted approach offers a scalable and practical pathway to reduce microbial exposure and advance public health protection, particularly in low-resource regions.

Existing national regulations serve as important empirical support for the validity of our proposed HDI-stratified *E. coli* discharge standards. For example, Kenya's zero-discharge policy represents an exceptionally stringent standard that frequently surpasses local operational capacities and imposes significant financial burdens (Fig. S14a,

Table S7). In contrast, China has enforced a 3 log MPN/100 mL fecal coliform limit since 2002 (GB 18918–2002), achieving over 93% disinfection coverage in municipal WWTPs (Figs. S9b and S14a, Table S7). This threshold aligns closely with our proposed standards for high-HDI countries (2.84 log MPN/100 mL for drinking water protection, 2.34 log MPN/100 mL for recreational water exposure), underscoring the feasibility of translating science-based targets into enforceable regulation when supported by sufficient infrastructure and governance (Figs. 4c,g, and S14a, Table S7).

Our findings highlight the value of globally tiered discharge standards, supported by targeted disinfection and adaptive policies, as a practical pathway for microbial risk reduction. The HDI-stratified framework offers a science-based yet flexible foundation, aligning with diverse regulatory capacities and infrastructure conditions worldwide. For lasting impact, discharge standards should be integrated into broader exposure-based regulations, such as the Clean Water Act [39] and the Bathing Water Directive [58], to comprehensively control microbial transmission risks. To enable feasible and scalable implementation across diverse regions, future efforts should emphasize phased targets, prioritized disinfection upgrades, and digital monitoring, alongside improved pathogen-specific data, enhanced surveillance, and stronger international collaboration to support health-based microbial risk management [59–61].

5. Conclusion

This study developed the first global *E. coli* emission database for 58,502 WWTPs across 188 countries and territories using a CatBoost machine learning model trained on actual data from 230 plants. By addressing a critical knowledge gap in microbial risk management, historically overshadowed by chemical concerns, this work reveals widespread and persistent microbial contamination. Results show that 78.3% of WWTPs discharge effluents exceeding 3 log MPN/100 mL, particularly in Africa, South America, and Southeast Asia. These contaminations are largely driven by limited disinfection coverage and the lack of enforceable microbial discharge standards. Effluent pollution from WWTPs threatens public health by degrading the microbial quality of adjacent drinking and recreational waters. To mitigate these risks, we propose health-based, HDI-stratified *E. coli* discharge thresholds, derived through reverse risk modeling and aligned with international guidelines. Simulation results show that coupling these standards with targeted disinfection upgrades could reduce global non-compliant WWTPs by 47% and 58% under drinking and recreational water protection, respectively, with particularly strong benefits for low- and middle-income countries. These findings underscore the urgent need for global action to establish standardized microbial discharge regulations, expand disinfection coverage, and strengthen international cooperation to reduce waterborne health risks.

CRediT authorship contribution statement

Wen Li: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Data curation, Conceptualization. **Qingbin Yuan:** Writing – review & editing, Validation, Supervision, Methodology, Investigation, Conceptualization. **Xiaolong Wang:** Writing – review & editing, Validation, Methodology, Investigation. **Wei Wang:** Validation, Methodology, Investigation, Data curation. **Xinda Wu:** Validation, Methodology, Investigation. **Yanxu Zhang:** Validation, Methodology. **Xiaoli Zhao:** Writing – review & editing, Validation, Funding acquisition. **Yi Luo:** Writing – review & editing, Validation, Supervision, Funding acquisition, Conceptualization. **Fengchang Wu:** Writing – review & editing, Validation, Supervision, Funding acquisition, Conceptualization.

Code availability

The original code used to develop the machine learning models in this study is publicly available on GitHub (https://github.com/liwensunny/WWTP_effluent_Ecoli_ML.git) for access and further scrutiny.

Data availability

All data utilized in this study are publicly accessible, as outlined in the Materials and methods section. National discharge regulations for WWTPs and surface water quality standards were systematically compiled from the FAOLEX database (Food and Agriculture Organization Legal Database, https://www.fao.org/faolex/country-profiles/en/) (Tables S7 and S8). The supplementary material includes a summary of the original data used for model development, along with their respective sources.

Declaration of competing interest

The authors declare no competing interests.

Acknowledgments

We thank the National Natural Science Foundation of China (42394153), the National Key R&D Program of China (2023YFC3205803), and the National Natural Science Foundation of China (42377436).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eehl.2025.100209.

References

[1] F. Silva, R. Al Bakain, G. Pradeau, M. Ben Braham, J. Prtorić, N. Ait-Mouheb, et al., Protection over restoration to ensure water sustainability, Commun. Earth Environ. 5 (2024) 638.
[2] World Health Organization, Progress on drinking-water and Sanitation: 2014 Update, World Health Organization, Geneva, 2014.
[3] E. Machdar, N.P. van der Steen, L. Raschid-Sally, P.N.L. Lens, Application of Quantitative Microbial Risk Assessment to analyze the public health risk from poor drinking water quality in a low income area in *Accra*, Ghana, Sci. Total Environ. 449 (2013) 134–142.
[4] G. Xu, T. Wang, Y. Wei, Y. Zhang, J. Chen, Fecal coliform distribution and health risk assessment in surface water in an urban-intensive catchment, J. Hydrol. 604 (2022) 127204.
[5] U. Ryan, K. Hill, D. Deere, Review of generic screening level assumptions for quantitative microbial risk assessment (QMRA) for estimating public health risks from Australian drinking water sources contaminated with *Cryptosporidium* by recreational activities, Water Res. 220 (2022) 118659.
[6] J.A. Soller, T. Bartrand, N.J. Ashbolt, J. Ravenscroft, T.J. Wade, Estimating the primary etiologic agents in recreational freshwaters impacted by human sources of faecal contamination, Water Res. 44 (16) (2010) 4736–4747.
[7] United Nations Environment Programme, A Snapshot of the World's Water Quality: towards a Global Assessment, United Nations Environment Programme, Nairobi, 2016.
[8] E.R. Jones, M.F.P. Bierkens, P.J.T.M. van Puijenbroek, L.P.H. van Beek, N. Wanders, E.H. Sutanudjaja, et al., Sub-Saharan Africa will increasingly become the dominant hotspot of surface water pollution, Nat. Water 1 (7) (2023) 602–613.
[9] E.R. Jones, M.F.P. Bierkens, N. Wanders, E.H. Sutanudjaja, L.P.H. van Beek, M.T. H. van Vliet, Current wastewater treatment targets are insufficient to protect surface water quality, Commun. Earth Environ. 3 (2022) 221.
[10] S. Itoh, L. Zhou, Effect of unboiled water consumption data on sensitivity analysis in quantitative microbial risk assessment, npj Clean Water 1 (2018) 18.
[11] J. Wolf, S. Hubbard, M. Brauer, A. Ambelu, B.F. Arnold, R. Bain, et al., Effectiveness of interventions to improve drinking water, sanitation, and handwashing with soap on risk of diarrhoeal disease in children in low-income and middle-income settings: a systematic review and meta-analysis, Lancet 400 (10345) (2022) 48–59.
[12] E. van Heijnsbergen, G. Niebaum, V. Lämmchen, A. Borneman, L. Hernández Leal, J. Klasmeier, et al., (antibiotic-resistant) *E. coli* in the Dutch–German vecht

catchment-monitoring and modeling, Environ. Sci. Technol. 56 (21) (2022) 15064–15073.
[13] J. Jang, H.-G. Hur, M.J. Sadowsky, M.N. Byappanahalli, T. Yan, S. Ishii, Environmental *Escherichia coli*: ecology and public health implications-a review, J. Appl. Microbiol. 123 (3) (2017) 570–581.
[14] T.A.T. Gomes, W.P. Elias, I.C.A. Scaletsky, B.E.C. Guth, J.F. Rodrigues, R.M. F. Piazza, et al., Diarrheagenic *Escherichia coli*, Braz. J. Microbiol. 47 (2016) 3–30.
[15] R.B.M. Marano, T. Fernandes, C.M. Manaia, O. Nunes, D. Morrison, T. U. Berendonk, et al., A global multinational survey of cefotaxime-resistant coliforms in urban wastewater treatment plants, Environ. Int. 144 (2020) 106035.
[16] World Health Organization, Guidelines for drinking-water Quality, World Health Organization, Geneva, 2002.
[17] United States Environmental Protection Agency, Recreational Water Quality Criteria, Office of Water, U.S. Environmental Protection Agency, Washington, D.C., 2012.
[18] European Union, Council Directive 76/160/EEC concerning the quality of bathing water, Off, J. Eur. Communities 19 (1976) 1–7.
[19] M. Levine-Clark, 'we the peoples:' the role of the united nations in the 21st century, J. Govern. Inf. 28 (5) (2001) 571–574.
[20] J. Foschi, A. Turolla, M. Antonelli, Soft sensor predictor of *E. coli* concentration based on conventional monitoring parameters for wastewater disinfection control, Water Res. 191 (2021) 116806.
[21] R. Zamudio, P. Boerlin, M.R. Mulvey, M. Haenni, R. Beyrouthy, J.-Y. Madec, et al., Global transmission of extended-spectrum cephalosporin resistance in *Escherichia coli* driven by epidemic plasmids, EBioMedicine 103 (2024) 105097.
[22] J.C. Semenza, Cascading risks of waterborne diseases from climate change, Nat. Immunol. 21 (5) (2020) 484–487.
[23] A. Dufour, J. Bartram, R. Bos, V. Gannon, Animal Waste, Water Quality and Human Health, IWA Publishing, London, 2012.
[24] D.M. Berendes, P.J. Yang, A. Lai, D. Hu, J. Brown, Estimation of global recoverable human and animal fecal biomass, Nat. Sustain. 1 (11) (2018) 679–685.
[25] S.R. Weirich, J. Silverstein, B. Rajagopalan, Effect of average flow and capacity utilization on effluent water quality from US municipal wastewater treatment facilities, Water Res. 45 (14) (2011) 4279–4286.
[26] H. Hashemi, M. Amin, A. Bovini, Y. Hung, A review on wastewater disinfection, Int. J. Environ. Health Eng. 2 (1) (2013) 22.
[27] H. Hersbach, B. Bell, P. Berrisford, G. Biavati, A. Horányi, J. Muñoz Sabater, et al., ERA5 monthly averaged data on single levels from 1940 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS) (2023).
[28] D. Krstajic, L.J. Buturovic, D.E. Leahy, S. Thomas, Cross-validation pitfalls when selecting and assessing regression and classification models, J. Cheminf. 6 (1) (2014) 10.
[29] X. Liu, Y. Nie, X.-L. Wu, Predicting microbial community compositions in wastewater treatment plants using artificial neural networks, Microbiome 11 (1) (2023) 93.
[30] J.T. Hancock, T.M. Khoshgoftaar, CatBoost for big data: an interdisciplinary review, J. Big Data 7 (1) (2020) 94.
[31] T. Chen, C. Guestrin, XGBoost: a scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, San Francisco California USA, 2016, pp. 785–794.
[32] L. Prokhorenkova, G. Gusev, A. Vorobev, A.V. Dorogush, A. Gulin, CatBoost: unbiased boosting with categorical features, Adv. Neural Inf. Process. Syst. 31 (2018).
[33] R.S. Hendriksen, P. Munk, P. Njage, B. van Bunnik, L. McNally, O. Lukjancenko, et al., Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage, Nat. Commun. 10 (1) (2019) 1124.
[34] T.C. Ekundayo, O.A. Ijabadeniyi, E.O. Igbinosa, A.I. Okoh, Using machine learning models to predict the effects of seasonal fluxes on *Plesiomonas shigelloides* population density, Environ. Pollut. 317 (2023) 120734.
[35] H. Ehalt Macedo, B. Lehner, J. Nicell, G. Grill, J. Li, A. Limtong, et al., Distribution and characteristics of wastewater treatment plants within the global river network, Earth Syst. Sci. Data 14 (2) (2022) 559–577.
[36] N.K. Arora, I. Mishra, Sustainable development goal 6: global water security, Environ. Sustain 5 (3) (2022) 271–275.
[37] I. Okoh Anthony, E. Odjadjare Emmanuel, O. Igbinosa Etinosa, N. Osode Augustina, Wastewater treatment plants as a source of microbial pathogens in receiving watersheds, Afr. J. Biotechnol. 6 (25) (2007) 2932–2944.
[38] M. Pulido-Velázquez, F.A. Ward, Comparison of water management institutions and approaches in the United States and Europe—What can we learn from each other? in: J.R. Ziolkowska, J.M. Peterson (Eds.), Competition for Water Resources Elsevier, Amsterdam, 2017, pp. 423–441.
[39] W.L. Andreen, Water quality today—has the Clean Water Act been a success, Ala. Law Rev. 55 (2003) 537–586.
[40] C. Tortajada, Contributions of recycled wastewater to clean water and sanitation Sustainable Development goals, npj Clean Water 3 (2020) 22.
[41] J.R. Bolton, C.A. Cotton, The Ultraviolet Disinfection Handbook, American Water Works Association, Denver, 2008.
[42] Black & Veatch Corporation, White'S Handbook of Chlorination and Alternative Disinfectants, fifth ed., John Wiley & Sons, Hoboken, 2011.
[43] T.A. Larsen, S. Hoffmann, C. Lüthi, B. Truffer, M. Maurer, Emerging solutions to the water challenges of an urbanizing world, Science 352 (6288) (2016) 928–933.
[44] P.R. Hunter, Climate change and waterborne and vector-borne disease, J. Appl. Microbiol. 94 (Suppl) (2003) 37S–46S.
[45] A. Hörman, Assessment of the Microbial Safety of Drinking Water Produced from Surface Water Under Field Conditions, University of Helsinki, Helsinki, 2005. Ph.D. dissertation.
[46] I.T. Miettinen, O. Zacheus, C.-H. von Bonsdorff, T. Vartiainen, Waterborne epidemics in Finland in 1998-1999, Water Sci. Technol. 43 (12) (2001) 67–71.
[47] N.A. Moreira, M. Bondelind, Safe drinking water and waterborne outbreaks, J. Water Health 15 (1) (2017) 83–96.
[48] H. Shuval, Estimating the global burden of thalassogenic diseases: human infectious diseases caused by wastewater pollution of the marine environment, J. Water Health 1 (2) (2003) 53–64.
[49] E.J. Dziuban, J.L. Liang, G.F. Craun, V. Hill, P.A. Yu, J. Painter, et al., Surveillance for waterborne disease and outbreaks associated with recreational water—United States, 2003–2004, MMWR Surveill, Summit 55 (2006) 1–30.
[50] K.L. Vanden Esschert, M.C. Mattioli, E.D. Hilborn, V.A. Roberts, A.T. Yu, K. Lamba, et al., Outbreaks associated with untreated recreational water: california, Maine, and Minnesota, 2018–2019, MMWR Morb. Mortal. Wkly. Rep. 69 (25) (2020) 781–783.
[51] R. Mitra, S.F. McGough, T. Chakraborti, C. Holmes, R. Copping, N. Hagenbuch, et al., Learning from data with structured missingness, Nat. Mach. Intell. 5 (1) (2023) 13–23.
[52] G. Joseph, Y.R. Hoo, Q. Wang, A. Bahuguna, L. Andres, Funding a Water-Secure Future: an Assessment of Global Public Spending, World Bank, Washington, D.C., 2024.
[53] S.I. Abou-Elela, M.M.H. El-Sayed, A.S. El-Gendy, E.M. Abou-Taleb, Comparative study of disinfection of secondary treated wastewater using chlorine, UV and ozone, J. Appl. Sci. Res. 8 (2012) 5190–5197.
[54] S. Tak, A. Kumar, Chlorination disinfection by-products and comparative cost analysis of chlorination and UV disinfection in sewage treatment plants: indian scenario, Environ. Sci. Pollut. Res. Int. 24 (34) (2017) 26269–26278.
[55] M. Ostoich, E. Aimo, R. Frate, M. Vazzoler, S. Stradella, P. Osti, Intergrated approach for microbiological impact assessment of public wastewater treatment plants, Chem. Ecol. 23 (1) (2007) 43–62.
[56] A. Simhon, V. Pileggi, C.A. Flemming, J.R. Bicudo, G. Lai, M. Manoharan, Enteric viruses in municipal wastewater effluent before and after disinfection with chlorine and ultraviolet light, J. Water Health 17 (5) (2019) 670–682.
[57] United States of America and State of Ohio, Second amendment to consent decree, United States district court for the Northern district of Ohio, Western Division, Toledo (2022) 1–16.
[58] C.R. Mansilha, C.A. Coelho, A.M. Heitor, J. Amado, J.P. Martins, P. Gameiro, Bathing waters: new directive, new standards, new quality approach, Mar. Pollut. Bull. 58 (10) (2009) 1562–1565.
[59] A. Mohanty, S.K. Mohanty, A.G. Mohapatra, Real-time monitoring and fault detection in AI-enhanced wastewater treatment systems, in: The AI Cleanse: Transforming Wastewater Treatment Through Artificial Intelligence: Harnessing Data-Driven Solutions, Springer Nature Switzerland AG, Cham, 2024, pp. 165–199.
[60] A. Elnaiem, O. Mohamed-Ahmed, A. Zumla, J. Mecaskey, N. Charron, M.F. Abakar, et al., Global and regional governance of one Health and implications for global health security, Lancet 401 (10377) (2023) 688–704.
[61] J. Pannu, M.J. Palmer, A. Cicero, D.A. Relman, M. Lipsitch, T. Inglesby, Strengthen oversight of risky research on pathogens, Science 378 (6625) (2022) 1170–1172.