4 (2026) 100034

Contents lists available at ScienceDirect

Journal of Hazardous Materials: Organics

journal homepage: www.sciencedirect.com/journal/journal-of-hazardous-materials-organics

# PAH–aliphatic hydrocarbon source disentanglement and ecological risk mapping in Ghanaian river sediments: National scale source apportionment across tropical basins

Ebenezer Aquisman Asare [a,*], Dickson Abdul-Wahab [b], Elsie Effah Kaufmann [c], Rafeah Wahi [d], Zainab Ngaini [d], Archibold Buah-Kwofie [e], Ibrahim Kwame Kwarteng [a], Crentsil Kofi Bempah [a], Joseph Richmond Fianko [b]

[a] Nuclear Chemistry and Environmental Research Centre, National Nuclear Research Institute (NNRI), Ghana Atomic Energy Commission (GAEC), Box LG 80, Legon-Accra, Ghana
[b] Department of Nuclear Science and Applications, School of Nuclear and Allied Sciences, University of Ghana, Atomic-Kwabenya, Accra, Ghana
[c] Department of Biomedical Engineering, University of Ghana, Legon-Accra, Ghana
[d] Department of Chemistry, Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, Kota Samarahan, Sarawak, Malaysia
[e] Nuclear Power Institute (NPI), Ghana Atomic Energy Commission (GAEC), Box LG 80, Legon-Accra, Ghana

ARTICLE INFO

*Keywords:*
Source disentanglement
Hydrocarbon Risk Surveillance Index
Unresolved complex mixture
Benzo[*a*]pyrene toxic equivalency
Mining urban pressure
Sediment capture

ABSTRACT

Hydrocarbon contamination in tropical river sediments is often assessed using polycyclic aromatic hydrocarbons (PAHs), although PAH only approaches can compress petroleum weathering, plant wax dilution and mixed anthropogenic source histories into broad pyrogenic classes. This study presents a national PAH–aliphatic hydrocarbon source disentanglement and Hydrocarbon Risk Surveillance Index (HRSI) framework for Ghanaian river sediments, based on 168 surface sediment records from 21 rivers across six operational basin systems, four reach positions and two hydrological seasons. For the first time at national river basin scale in Ghana, EPA-16 PAHs are integrated with C10–C35 *n*-alkanes, pristane/phytane, carbon preference index, terrigenous/aquatic ratio, unresolved complex mixture fractions, exploratory PCA/HCA, constrained non-negative latent source classification and HRSI screening to separate petroleum weathering signals from combustion source signals. Basin scale hydrocarbon burden followed the surveillance priority order Ankobra > Pra > Coastal > Volta > Tano > Bia. PAH diagnostic ratios classified most samples as pyrogenic dominant, whereas aliphatic biomarkers and latent source fractions identified additional weathered-petroleum, mining-urban, traffic-combustion, biomass-combustion and terrestrial-biogenic signatures. Risk was spatially compressed: 114 samples were low risk, 40 were moderate, 12 were high and 2 were severe, with severe dry season records at the Densu estuary and Ankobra midstream. The results identify sediment capture, especially in depositional or low gradient reaches, as a critical process by which petroleum weathering and combustion signals are converted into local high risk hotspots. Mining pressure, urban traffic pressure and depositional reach position were positively associated with hydrocarbon burden and HRSI class, but the observational design and collinearity among urban land use, road density and population density require interpretation as associations rather than causal effects. The study provides a testable basis for basin prioritised hydrocarbon surveillance in Ghana.

## 1. Introduction

Hydrocarbon contamination in riverine and estuarine sediments remains a persistent environmental concern because sediments act simultaneously as sinks, reaction media, and historical archives of petroleum, combustion, urban, industrial, transport, mining, and biomass burning inputs. Polycyclic aromatic hydrocarbons (PAHs) are particularly important because many are hydrophobic, particle reactive, environmentally persistent, and toxicologically relevant, while petroleum derived aliphatic hydrocarbons, including *n*-alkanes, isoprenoids and unresolved complex mixtures, provide additional molecular evidence for petroleum residues, weathering, biodegradation and organic matter

* Corresponding author.
*E-mail address:* aquisman1989@gmail.com (E.A. Asare).

https://doi.org/10.1016/j.hazmo.2026.100034
Received 18 May 2026; Received in revised form 7 June 2026; Accepted 13 June 2026
Available online 17 June 2026
3051-0597/© 2026 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

sources. Recent sediment studies continue to show that PAH distributions alone are commonly interpreted through diagnostic ratios, multivariate statistics and ecological risk indices, but source discrimination becomes stronger when PAHs are integrated with aliphatic markers and spatial risk mapping [1–4]. For example, recent work on subtropical and tropical sediments has combined PAH occurrence, source apportionment and ecological risk assessment to distinguish combustion dominated inputs from local petroleum contributions and to identify priority monitoring zones.

In Ghana, the relevance of hydrocarbon fingerprinting is heightened by the structure of the national drainage network and the diversity of anthropogenic pressures intersecting the major basins. The Volta, Southwestern and Coastal basin systems connect mining belts, urban centres, transport corridors, agricultural landscapes, wetlands and coastal depositional zones, creating multiple pathways for petroleum residues, combustion derived PAHs, biomass burning products and industrial hydrocarbons to accumulate in aquatic sediments. Existing Ghanaian evidence has documented PAHs in the Densu River basin and Fosu Lagoon, while recent offshore work has combined aliphatic and PAH fingerprints with probabilistic source and risk assessment [5–7]. These studies provide important local, lagoonal, coastal or offshore evidence, but they do not provide a nationally integrated river sediment dataset that combines EPA−16 PAHs, aliphatic biomarkers, UCM structure, source classification and screening level risk across major Ghanaian basins. Consequently, three critical questions remain unresolved: whether Ghanaian river sediments exhibit distinct national PAH-aliphatic hydrocarbon fingerprints; whether PAH diagnostic ratios alone are sufficient for source discrimination under mixed tropical catchment conditions; and which basins or reaches should be prioritised for long term ecological risk surveillance.

This paper therefore moves beyond conventional contaminant measurement toward national molecular source screening. PAH diagnostic ratios remain useful for first order discrimination between petrogenic and pyrogenic sources, but they may be ambiguous where petroleum residues, biomass burning, traffic emissions, industrial combustion, sediment transport, weathering and mixed source deposition occur together. Recent studies reinforce this limitation by increasingly pairing diagnostic ratios with principal component analysis, receptor style source apportionment, ecological risk quotients, toxic equivalent assessment and spatial mapping [3,4,8–10]. Aliphatic biomarkers also require cautious interpretation because CPI, TAR, pristane/phytane and UCM can be modified by source facies, biodegradation, organic matter mixing, sediment texture and hydrodynamic sorting. For this reason, the present study treats both PAH ratios and aliphatic indices as complementary evidence layers rather than definitive source fingerprints. The 2024 study on PAHs and *n*-alkanes in polluted estuarine water and sediments demonstrates the value of integrating aromatic and aliphatic hydrocarbons for source and ecological risk interpretation [3]. Similarly, a West African mangrove sediment study from Ifiekporo Creek, Nigeria, used GC–MS, isomeric ratios and multivariate statistics to infer mixed pyrogenic and petrogenic PAH inputs, illustrating the regional relevance of combined molecular and statistical source screening for sediment management [2]. The main innovation is not the general observation that PAH only source diagnosis can be limited. Rather, the contribution is the first national level PAH-aliphatic source disentanglement assessment of Ghanaian river sediments, in which petroleum-weathering/UCM signals and combustion derived PAH signals are evaluated together across competing basins, reach positions and seasons. This formulation also identifies sediment capture in lower depositional or transitional reaches as a process that converts catchment inputs into spatially compressed risk hotspots.

Positive matrix factorisation (PMF) is also important for hydrocarbon source apportionment because it estimates non-negative source profiles and contributions while incorporating uncertainty information and residual diagnostics, thereby reducing subjectivity compared with diagnostic ratio only interpretation [17,27]. In this study, PMF is discussed as a benchmark receptor modelling framework, while the implemented constrained non-negative latent fraction model is used as a screening analogue suitable for the dual PAH-aliphatic matrix and the available sample size. Accordingly, this study aims to construct a national PAH-aliphatic hydrocarbon fingerprint atlas for Ghanaian river sediments; identify basin, reach and season dependent source patterns; compare diagnostic ratio classifications with latent source classifications; map screening level ecological risk classes across major river systems; and propose a practical monitoring template for Ghana and comparable tropical basins. Four testable expectations guided the analysis: higher pressure mining and urban basins would show higher ΣPAH16, BaP-TEQ, total aliphatic burden and UCM fractions than lower pressure basins; PAH diagnostic ratios would classify many samples as pyrogenic but would not fully separate traffic, biomass, mining urban and petroleum weathering signals; downstream and depositional reaches would show higher hydrocarbon burden than upper/reference reaches; and dry season samples would show stronger concentration of sediment associated hydrocarbons than wet season samples.

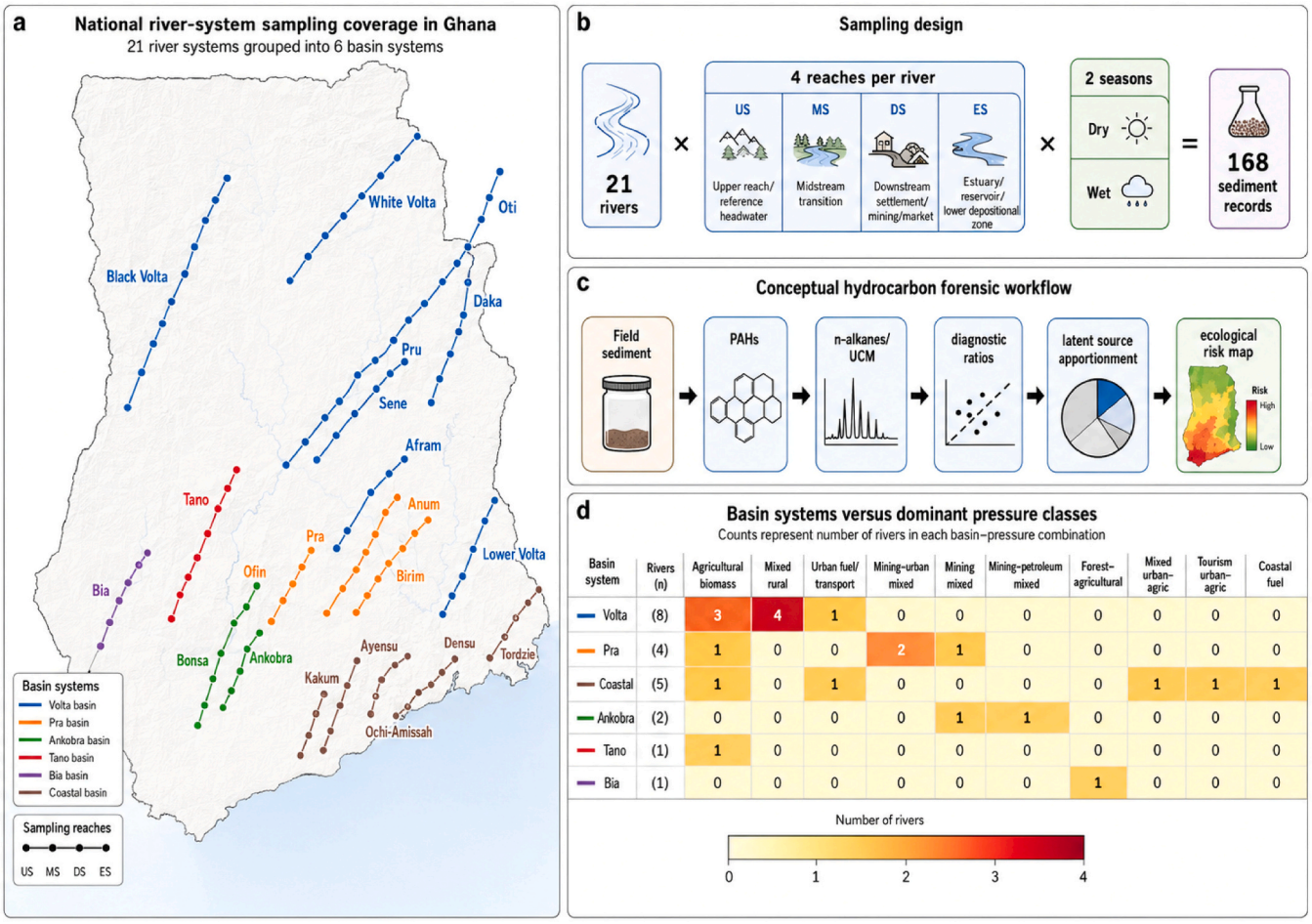
## 2. Materials and methods

### 2.1. Study area and national sampling design

This study was designed as a national sediment hydrocarbon source disentanglement survey across Ghana's major river systems. Ghana's drainage network is commonly described within three principal outlet domains; the Volta, Southwestern and Coastal domains but the present study used six operational basin systems to preserve hydrological and management meaning at the scale of sampling and interpretation. The Bia, Tano, Ankobra and Pra were treated separately because they are hydrologically independent Southwestern basins with different mining, urban, agricultural and depositional contexts; grouping them into a single Southwestern class would have hidden the strong Ankobra-Pra mining-urban signal and the lower burden Bia/Tano endmembers. The Coastal system was retained as an operational class for short, independent coastal rivers such as the Densu, Ayensu and Kakum because they share a coastal outlet setting, short catchment length, urban/traffic influence and estuarine depositional behaviour, although they are not interpreted as one hydrologically connected river basin. This rationale is reflected in Fig. 1d and the Supplementary Information. The national network therefore comprised 21 rivers, six operational basin systems and four reach classes: upper/reference reaches, mid-catchment agricultural or mining influenced reaches, urban-industrial or transport corridor reaches, and lower depositional/estuarine reaches. Sampling was conducted in both dry and wet seasons to capture seasonal contrasts in runoff mobilisation, combustion derived atmospheric deposition, urban wash-off, sediment resuspension and depositional focusing. Station level pressure covariates were extracted within a common 1 km buffer around each sampling point because the study sought to test local near channel pressure on surface sediment fingerprints rather than to model whole catchment pollutant loads. The 1 km buffer captures proximal road runoff, riparian urbanisation, mining disturbance, bank side fuel handling and local land use inputs that can directly influence deposited surface sediment. A catchment scale analysis would be valuable for flux estimation, but it was not used as the primary covariate framework here because catchment sizes, flow routing and upstream retention differ strongly among the 21 rivers and would mix regional background pressure with the local depositional signal targeted by this sediment survey. These covariates were therefore used as local pressure indicators for statistical association, not as direct measurements of basin scale hydrocarbon discharge rates.

### 2.2. Sediment collection and preservation

At each station, surface sediment was collected from the biologically and geochemically active upper layer of the bed sediment,




| Basin system | Rivers (n) | Agricultural biomass | Mixed rural | Urban fuel/ transport | Mining-urban mixed | Mining mixed | Mining-petroleum mixed | Forest-agricultural | Mixed urban-agric | Tourism urban-agric | Coastal fuel |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Volta | (8) | 3 | 4 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Pra | (4) | 1 | 0 | 0 | 2 | 1 | 0 | 0 | 0 | 0 | 0 |
| Coastal | (5) | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 |
| Ankobra | (2) | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| Tano | (1) | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Bia | (1) | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |

**Fig. 1.** National sampling architecture and hydrocarbon forensic design. National scale sampling design used for hydrocarbon source disentanglement of Ghanaian river sediments. (a) Spatial distribution of the 21 river systems grouped into six operational basin systems: Volta, Pra, Ankobra, Tano, Bia and Coastal. Bia, Tano, Ankobra and Pra are treated separately because they are hydrologically independent Southwestern basins, whereas short independent coastal rivers are grouped operationally because they share coastal outlet and depositional behaviour. (b) Sampling architecture showing four reach positions and two hydrological seasons. (c) Conceptual workflow linking field sediment, PAHs, n-alkanes/UCM, diagnostic ratios, latent source classification and HRSI. (d) Basin systems versus dominant pressure classes used to justify the operational grouping.

corresponding to approximately 0–5 cm depth. Where channel morphology permitted, five to seven sub-samples were taken across a depositional micro-reach using pre-cleaned stainless steel tools and composited to obtain a representative sample. Coarse debris, leaves, shells, visible plastics and stones were removed in the field without disturbing the fine sediment fraction. Field duplicates, equipment blanks and procedural field blanks were included to assess sampling precision and contamination control. All sediment samples were transferred into pre-combusted amber glass containers with solvent rinsed aluminium foil lined caps. Samples were stored on ice in the field, transported to the laboratory under cooled conditions, and kept at $\leq 4$ °C for short-term storage or frozen at −20 °C until processing. Before extraction, sediments were freeze dried or oven dried at low temperature, homogenised, sieved, and stored in solvent rinsed glassware. Moisture content, grain size and total organic carbon were determined for normalisation and interpretation because hydrocarbon accumulation in sediments is strongly controlled by fine particle abundance and organic-matter content.

### 2.3. Hydrocarbon extraction, fractionation and instrumental analysis

Hydrocarbons were extracted from dried and homogenised sediment aliquots using one harmonised primary protocol to avoid method driven differences among basins. Approximately 10 g of dry sediment was spiked with deuterated PAH surrogates and aliphatic recovery standards, mixed with pre-combusted anhydrous sodium sulphate, and extracted by pressurised fluid extraction following USEPA Method 3545 A principles [25]. The extraction solvent was *n*-hexane:dichloromethane (1:1, v/v), with two static extraction cycles, nitrogen purge and solvent-rinsed collection vessels. Where repeat or validation extracts were produced by ultrasonic or Soxhlet extraction, those runs were used only for QA/QC comparability checks and were not mixed with the primary concentration matrix unless they passed the same blank, recovery and duplicate precision criteria. This clarification is important because different extraction routes can alter recoveries and apparent UCM abundance. The crude extract was concentrated under a gentle nitrogen stream and cleaned by column chromatography using activated silica gel and alumina. The aliphatic fraction was eluted first with *n*-hexane, and the aromatic fraction was then eluted with dichloromethane:n-hexane. Elemental sulphur, where present, was removed using activated copper. Quality assurance and quality control covered the full workflow from field collection to instrumental reporting. Field duplicates, equipment blanks and field blanks evaluated collection precision and external contamination; procedural and method blanks evaluated laboratory contamination; matrix spikes, surrogate standards and internal standards evaluated recovery and matrix effects; continuing calibration checks and laboratory reference materials evaluated instrumental stability. Multi-point calibration was required for all

target PAHs and *n*-alkanes, with quantification accepted only when calibration, blank and continuing check evidence were internally consistent. Limits of detection and quantification were defined from low level standard/blank variability and signal-to-noise verification, with LOQ treated as the lowest level suitable for quantitative reporting. The expanded QA/QC verification matrix in Supplementary Table 9 f now records the calibration, blank, recovery, precision and reporting status checks used to avoid unsupported numerical claims. QA/QC outcomes were retained in the dataset rather than hidden: 168 of 168 expected sediment records were obtained, 156 records were graded A and 12 were graded B, 165 records passed method blank review, mean surrogate recovery was 86.8% (median 86.7%; range 66.7–107.1%), mean duplicate RPD was 12.6% (median 11.7%; range 1.8–25.9%), PAH detection limits were 0.01–0.05 ng $g^{-1}$ dry weight and n-alkane detection limits were 0.02–0.10 ng $g^{-1}$ dry weight (Supplementary Table 9). Method-blank review flags and Grade B records were retained transparently, so downstream source and risk interpretation reflects residual analytical uncertainty rather than presenting a falsely perfect dataset. This full process reporting structure follows recent contaminant monitoring practice in which blanks, detection limits, calibration performance, matrix effects, recovery and precision are reported together rather than as isolated QA/QC statements [27,28].

The aromatic fraction was analysed for the 16 priority PAHs commonly used in environmental monitoring: naphthalene, acenaphthylene, acenaphthene, fluorene, phenanthrene, anthracene, fluoranthene, pyrene, benz[*a*]anthracene, chrysene, benzo[*b*]fluoranthene, benzo[*k*]fluoranthene, benzo[*a*]pyrene, indeno[1,2,3-cd]pyrene, dibenz[*a*,*h*]anthracene and benzo[*g*,*h*,*i*]perylene. Quantification was performed by gas chromatography-mass spectrometry in selected-ion monitoring mode using multi-point calibration and internal standard correction [26]. The complete EPA−16 PAH concentration matrix is provided in Supplementary Table 2. The aliphatic fraction was analysed for resolved C10–C35 *n*-alkanes and isoprenoids, especially pristane and phytane. Resolved *n*-alkanes and isoprenoids were quantified against homologous standards, while the unresolved complex mixture was quantified by integrating the chromatographic hump above the baseline after subtraction of resolved peak areas. The UCM fraction was calculated as:

$$\text{UCM fraction} = \frac{\text{UCM area}}{\text{UCM area} + \text{resolved aliphatic hydrocarbon area}} \quad (1)$$

A high UCM fraction was interpreted as evidence consistent with weathered, biodegraded or chronically introduced petroleum residues, following petroleum geochemical interpretation principles [7,11,12]. Because UCM may also be influenced by matrix complexity and degradation state, it was not used alone as proof of petroleum input. The complete aliphatic, isoprenoid and UCM dataset is provided in Supplementary Table 3.

### 2.4. PAH diagnostic ratios

PAH diagnostic ratios were calculated to provide first-order source interpretation. The following ratios were used:

$$\frac{ANT}{ANT + PHE}$$

$$\frac{FLT}{FLT + PYR}$$

$$\frac{BaA}{BaA + CHR}$$

$$\frac{IP}{IP + BghiP}$$

$$\frac{LMW\ PAHs}{HMW\ PAHs} \quad (2)$$

where ANT = anthracene, PHE = phenanthrene, FLT = fluoranthene, PYR = pyrene, BaA = benz[*a*]anthracene, CHR = chrysene, IP = indeno[1,2,3-cd]pyrene and BghiP = benzo[*g*,*h*,*i*]perylene. Low molecular weight PAHs were defined as two to three ring compounds, while high molecular weight PAHs were defined as four to six ring compounds. The diagnostic ratios were interpreted using established petrogenic-pyrogenic thresholds but not treated as definitive source evidence because PAH ratios may be altered by weathering, photodegradation, sediment sorting, combustion temperature variability and source mixing [13,14]. Ratio based interpretation was therefore used as a screening layer and subsequently compared with latent source classifications.

### 2.5. Aliphatic biomarker indices

Aliphatic hydrocarbon indices were calculated to provide an independent petroleum and organic matter source domain. The pristane/phytane ratio was used to evaluate petroleum maturity, redox history and biodegradation influence, although it was interpreted cautiously because Pr/Ph may vary with source facies and alteration history [12]. The carbon preference index for higher plant wax hydrocarbons was calculated over the $C_{25}$–$C_{33}$ range as:

$$CPI_{25-33} = 0.5\left[\frac{C_{25} + C_{27} + C_{29} + C_{31} + C_{33}}{C_{24} + C_{26} + C_{28} + C_{30} + C_{32}} \middle| \frac{C_{25} + C_{27} + C_{29} + C_{31} + C_{33}}{C_{26} + C_{28} + C_{30} + C_{32} + C_{34}}\right] \quad (3)$$

Values substantially above unity were interpreted as stronger odd carbon number dominance and greater higher-plant wax contribution, whereas values approaching unity were interpreted as more petroleum-like, thermally mature or degraded signatures [12,15].

The terrigenous/aquatic ratio was calculated as:

$$TAR = \frac{C_{27} + C_{29} + C_{31}}{C_{15} + C_{17} + C_{19}} \quad (4)$$

where high TAR values indicate greater higher-plant wax input and low values indicate stronger aquatic, algal, microbial or petroleum related lower chain hydrocarbon influence [16]. Additional indicators included the ratio of low to high molecular weight *n*-alkanes and the UCM fraction. These aliphatic indices were interpreted jointly and cautiously because plant wax input, petroleum maturity, biodegradation, sediment sorting and organic carbon dilution can all modify the same biomarker ratios.

### 2.6. Dual-domain source apportionment framework

Source attribution was performed using a dual domain source screening framework integrating aromatic and aliphatic molecular evidence. The aromatic domain included individual PAHs, ring-class distributions and diagnostic ratios. The aliphatic domain included C10–C35 n-alkane homologues, Pr/Ph, CPI, TAR, UCM fraction and LMW/HMW *n*-alkane structure. Before multivariate analysis, concentrations below detection limits were replaced by one-half of the compound specific detection limit, variables were log10(x + 1) transformed where concentration distributions were highly skewed, and variables were autoscaled before multivariate classification. Observed molecular evidence and model-derived source labels were kept analytically separate during interpretation.

Principal component analysis and hierarchical cluster analysis were used only as exploratory tools to identify basin groupings, co-varying hydrocarbon assemblages and reach specific fingerprints. Source fractions were then estimated using a constrained non-negative matrix factorisation latent-fraction model. Candidate solutions containing three to seven components were compared using residual structure,

interpretability, bootstrap stability and agreement with independent basin and pressure metadata. The five-component solution was retained because it gave interpretable and stable classes corresponding to biomass combustion, traffic combustion, mining–urban mixed input, terrestrial-biogenic input and petrogenic input. The source classes are therefore reported as screening level latent signatures rather than definitive proof of single physical sources. This approach builds on receptor-modelling principles developed for environmental source apportionment, including PCA/HCA, PMF and NMF frameworks [17–19], and extends recent Ghanaian hydrocarbon-forensic work applying Bayesian source apportionment to aliphatic and PAH contamination in offshore sediments [7]. PMF was not selected as the primary fraction engine because the study objective was screening level latent classification from a combined PAH-aliphatic matrix rather than definitive quantitative receptor apportionment with a fully developed site-specific uncertainty matrix. Nevertheless, PMF principles guided the non-negativity constraint, component interpretability checks, residual review and bootstrap style stability screening. The model output is therefore presented as source class surveillance evidence.

### 2.7. Ecological risk screening

Ecological risk was evaluated using the Hydrocarbon Risk Surveillance Index (HRSI), a composite surveillance metric designed to rank sites and basins for monitoring priority. HRSI is explicitly a screening tool, not a regulatory threshold, toxicity standard or site-specific ecological effect prediction. It combines three evidence components that represent different hazard dimensions: ΣPAH16 burden (overall priority PAH load), BaP-equivalent toxicity (toxic potency of the PAH mixture) and UCM/petroleum weathering burden (chronic or degraded petroleum residue not captured fully by EPA−16 PAHs). Each component was normalised to its national 90th percentile before averaging so that the index retained sensitivity to extreme hotspots without allowing the two most extreme samples to define the entire national scale. The toxicological relevance of PAHs and benzo[*a*]pyrene supports the use of a BaP-equivalent component in the index [23,24].

The ΣPAH16 burden represented the total concentration of the 16 priority PAHs. BaP-equivalent toxicity was calculated as:

$$BaP_{eq} = \sum C_i \times TEF_i \quad (5)$$

where Ci is the concentration of each PAH and TEFi is its benzo[*a*]pyrene toxic equivalency factor. This approach follows the toxic equivalency logic commonly used to express the relative potency of PAH mixtures [20]. The HRSI for each sample was calculated as the arithmetic mean of the P90-normalised ΣPAH16 burden, P90-normalised BaP-TEQ and P90-normalised UCM burden. The UCM/petroleum weathering component captured chronic or degraded petroleum residues that may not be adequately represented by priority PAH concentrations alone. Where appropriate, PAH concentrations were additionally compared with sediment quality guideline values such as ERL/ERM or TEL/PEL benchmarks to support qualitative risk interpretation [21,22]. Compound specific toxic equivalency factors and compound metadata are provided in Supplementary Table 10. HRSI class thresholds were empirical surveillance bands on the P90-normalised index scale rather than externally defined toxicity thresholds: Low $< 0.25$, Moderate $= 0.25$ to $< 0.75$, High $= 0.75$ to $< 2.00$ and Severe $\geq 2.00$. The thresholds were selected to separate background or near background records, elevated monitoring records, high priority records and extreme outlier records in the national distribution. To assess whether HRSI behaved consistently with established sediment quality tools, index values and classes were cross checked against ERL/ERM and TEL/PEL outcomes using guideline exceedance counts, maximum guideline quotients and rank correlations. Sensitivity analysis recalculated HRSI after doubling the BaP-TEQ weight, removing the UCM component and excluding the two severe high leverage samples. Basin order, severe site identity and high/severe class membership were then compared with the base model to determine whether surveillance priorities were robust to weighting choices.

### 2.8. Statistical and geospatial analysis

Statistical analysis was conducted at basin, river, reach and seasonal scales. Because the dataset has nested structure, with samples grouped by river, basin, reach and season, basin rankings were interpreted as screening comparisons rather than as fully independent observations. Distributional assumptions were checked before formal comparisons; non-normal and heteroscedastic concentration variables were interpreted using $\log10(y + 1)$ transformations, rank-based tests and robust summaries. Seasonal contrasts were evaluated using paired station logic where dry and wet records existed for the same river-reach cell. Longitudinal gradients were assessed across reach classes to identify upstream-downstream enrichment, urban amplification, mining-associated signatures and depositional accumulation. Source risk coupling was examined using Spearman rank correlations, transparent bivariate OLS regressions and collinearity robust sensitivity models. Response variables included $\log10(y + 1)$-transformed ΣPAH16 and BaP-TEQ, raw UCM fraction and HRSI. Predictor variables included mining pressure, agriculture, forest cover and an urban traffic pressure index calculated as the mean of z-standardised urban land use, road density and population density. The index was used because the three original variables were strongly intercorrelated and VIF diagnostics showed that their coefficients could not be interpreted as independent effects. Ridge regression and partial least squares regression were used as sensitivity checks; conclusions are therefore expressed as associations rather than causal effects. To separate source input effects from depositional enrichment, partial correlations were calculated among ΣPAH16, TOC and silt-clay fraction. The analysis evaluated whether ΣPAH16 remained associated with mining or urban traffic pressure after controlling for TOC and particle size, and whether TOC or silt-clay fraction remained associated with ΣPAH16 after controlling for anthropogenic pressure.

## 3. Results

### 3.1. National hydrocarbon atlas reveals strong basin structuring

The national dataset produced a complete hydrocarbon fingerprinting matrix comprising 168 surface sediment records from 21 river systems, grouped into six basin systems and sampled across four reach positions and two hydrological seasons (Fig. 1; Table 1; Supplementary Table 1; Appendix Fig. 1). The spatial design confirms that the dataset has balanced reach and seasonal representation across the Ghanaian river continuum. Because basin sample sizes differed, with Volta and Coastal represented by more records than Tano and Bia, basin rankings are interpreted with attention to uncertainty, high leverage samples and nested river structure. The basin ranked hydrocarbon atlas shows that contamination is not randomly distributed across Ghanaian river systems. Instead, ΣPAH16 burden, BaP-equivalent toxicity, total aliphatic hydrocarbons and HRSI values define the basin-priority order Ankobra > Pra > Coastal > Volta > Tano > Bia (Fig. 2; Table 2; Supplementary Table 7a). Ankobra was the highest-priority basin, with a mean ΣPAH16 of 3267.3 ng $g^{-1}$ , mean BaP-TEQ of 258.8 ng $g^{-1}$ , mean total aliphatics of 11,706.8 ng $g^{-1}$ , and six high/severe records. Pra ranked second, with mean ΣPAH16 of 2183.4 ng $g^{-1}$ , mean BaP-TEQ of 191.7 ng $g^{-1}$ , mean total aliphatics of 6511.0 ng $g^{-1}$ , and six high/severe records. Coastal systems showed lower mean ΣPAH16 than Ankobra and Pra but contained the highest single national outlier, whereas Volta showed broad spatial coverage but mostly low-risk conditions with isolated hotspots. Tano and Bia formed the low burden endmember systems, with Bia showing the strongest low risk and terrestrial-biogenic signal. This ordering provides evidence for basin specific hydrocarbon

**Table 1**
National study design and dataset architecture.

| Component | Dataset content |
|---|---|
| Spatial coverage | 21 river systems grouped into 6 basin systems: Volta, Pra, Ankobra, Tano, Bia and Coastal |
| Samples | 168 surface sediment records |
| Design | 4 reach positions × 2 seasons |
| Reach positions | Upper/reference; midstream transition; downstream settlement/mining/market; estuary/reservoir/lower depositional reach |
| Seasons | Dry and wet seasons |
| Matrix | Surface sediment, 0–5 cm |
| PAH variables | EPA−16 PAHs, ΣPAH16, LMW PAHs, HMW PAHs and BaP-TEQ |
| Aliphatic variables | n-alkanes C10–C35, pristane, phytane and unresolved complex mixture |
| Biomarker indices | Pristane/phytane, CPI C25–C33, TAR and UCM fraction |
| Source tools | PAH diagnostic ratios; aliphatic biomarker indices; exploratory PCA/HCA; constrained NMF latent fractions |
| Risk tools | Hydrocarbon Risk Surveillance Index (HRSI) with fixed Low, Moderate, High and Severe surveillance classes |
| Environmental covariates | Urban land use, mining pressure, agriculture, forest cover, road density, population density, TOC, black carbon and sediment texture |

**Note:** PAH, polycyclic aromatic hydrocarbon; ΣPAH16, sum of the 16 EPA priority PAHs; BaP-TEQ, benzo[*a*]pyrene toxic equivalent; CPI, carbon preference index; TAR, terrigenous/aquatic ratio; UCM, unresolved complex mixture; TOC, total organic carbon.

syndromes in Ghana. Ankobra and Pra represent mining-urban mixed and pyrogenic high burden systems; Coastal rivers represent traffic/urban-combustion hotspot systems; Volta represents a broad but mostly low risk biomass combustion background with isolated urban or petroleum influenced anomalies; Tano represents a low burden biomass combustion system; and Bia represents a low burden terrestrial biogenic system (Fig. 2; Supplementary Table 7a).

### 3.2. Ecological risk is compressed into a small number of intense hotspots

The national HRSI distribution was highly compressed. Of the 168 sediment records, 114 were classified as low HRSI, 40 as moderate, 12 as high, and only 2 as severe (Fig. 6; Supplementary Table 6). Thus, 91.7% of samples were low or moderate, while only 8.3% were high or severe. However, the high/severe minority controlled the national hazard structure because they carried disproportionate PAH burden, toxicity-equivalent load and aliphatic petroleum-weathering signal. The two severe records define the national hotspot endmembers. The most extreme sample was GH-HCF−16-ES-D from the Densu River, Coastal basin, collected from the estuary/lower depositional reach during the dry season. It recorded ΣPAH16 = 15,895.6 ng $g^{-1}$, BaP-TEQ = 1396.6 ng $g^{-1}$, total aliphatics = 50,086.28 ng $g^{-1}$, UCM = 15,585.92 ng $g^{-1}$, and a dominant latent source of traffic combustion. The second severe record was GH-HCF−12-MS-D from the Ankobra River, collected from the midstream reach during the dry season. It recorded ΣPAH16 = 10,378.74 ng $g^{-1}$, BaP-TEQ = 788.66 ng $g^{-1}$, total aliphatics = 46,734.85 ng $g^{-1}$, UCM = 12,476.87 ng $g^{-1}$, and a dominant latent source of mining–urban mixed input (Supplementary Table 2; Supplementary Table 3; Supplementary Table 6; Supplementary Table 10; Appendix Fig. 8). Sensitivity checking showed that the basin order was not solely produced by these two high-leverage samples: after excluding them, mean ΣPAH16 remained highest in Ankobra (2793.2 ng $g^{-1}$), followed by Pra (2183.4 ng $g^{-1}$), Coastal systems (991.3 ng $g^{-1}$), Volta, Tano and Bia. This pattern indicates risk compression: most of the Ghanaian river sediment network carries low to moderate HRSI surveillance risk, but the upper tail of national hazard is controlled by a small number of basin-reach-season combinations. The severe hotspots occur in dry season depositional or transitional settings and are associated with strong anthropogenic source signatures rather than diffuse basin wide contamination. HRSI validation against sediment quality guideline outcomes produced the same surveillance interpretation as the base risk map. Samples in the high/severe HRSI classes were also the samples with the greatest ERL/ERM or TEL/PEL exceedance burden and the highest maximum guideline quotients, whereas low HRSI samples were generally below or close to lower effect range values. Doubling the BaP-TEQ weight increased emphasis on toxicity-dominant PAH mixtures but retained the basin-priority order Ankobra > Pra > Coastal > Volta > Tano > Bia and preserved the two severe records. Removing the UCM component reduced petroleum-weathering sensitivity in the Densu and Ankobra hotspots but did not remove them from priority status. HRSI is therefore robust for surveillance ranking.

### 3.3. PAH diagnostic ratios indicate pyrogenic dominance but under-resolve source mixtures

PAH diagnostic ratio classification showed overwhelming pyrogenic dominance. Across the 168 sediment records, 134 samples were classified as pyrogenic dominant, while 34 samples were classified as mixed petro-pyrogenic (Fig. 3; Supplementary Table 4). The diagnostic space therefore suggests that combustion derived inputs dominate Ghanaian river sediment PAH composition at national scale. The molecular PAH profiles support this interpretation. Normalised EPA−16 PAH fingerprints and LMW/HMW relationships show widespread enrichment of higher-molecular-weight PAHs, especially fluoranthene, pyrene and other four to six ring compounds, consistent with combustion related sources (Fig. 3a,b; Appendix Fig. 2). Diagnostic cross plots based on ANT/(ANT+PHE), FLT/(FLT+PYR), BaA/(BaA+CHR) and IP/(IP+BghiP) separate pyrogenic and mixed petro-pyrogenic domains, but they do not fully resolve whether combustion signatures arise from biomass burning, traffic emissions, mining associated fuel use, urban waste burning or petroleum-weathered residues (Fig. 3c–e; Appendix Fig. 4). Thus, PAH ratios provide a useful first order source screen, but they make Ghana's river systems appear more uniformly pyrogenic than they are. Their main limitation is source compression: diverse anthropogenic and natural inputs are collapsed into broad combustion dominant classes. This limitation becomes evident when the PAH results are compared with aliphatic biomarkers and latent source fractions.

### 3.4. Aliphatic biomarkers expose hydrocarbon-source complexity hidden by PAHs

Aliphatic biomarker evidence revealed source complexity that was not resolved by PAH diagnostic ratios alone. The C10–C35 *n*-alkane distributions, pristane/phytane ratios, CPI C25–C33, TAR values and UCM fractions distinguished multiple hydrocarbon domains, including weathered petroleum residues, mining-urban mixtures, traffic/urban fuel inputs, biomass-combustion backgrounds and terrestrial plant wax contributions (Fig. 4; Supplementary Table 3; Appendix Fig. 3). The aliphatic patterns were especially important in the high pressure basins. Ankobra and Pra showed elevated total aliphatic hydrocarbon burdens and stronger UCM contributions, consistent with complex anthropogenic hydrocarbon mixtures and weathered petroleum residues. Coastal systems showed strong traffic/urban fuel behaviour, particularly in the Densu severe record, where resolved *n*-alkanes were superimposed on a broad UCM hump. By contrast, Tano and Bia showed much lower total aliphatic burdens and stronger low risk terrestrial-biogenic or biomass associated signatures. The biomarker indices therefore changed the source interpretation from a simple PAH-based statement of "pyrogenic dominance" to a more differentiated screening classification of combustion type, petroleum weathering, terrestrial organic input and mixed mining-urban hydrocarbon loading. This is one of the central advances of the dataset: PAHs identify broad combustion pressure, whereas aliphatic biomarkers identify hydrocarbon processing history and non-aromatic petroleum/biogenic components that help explain the full

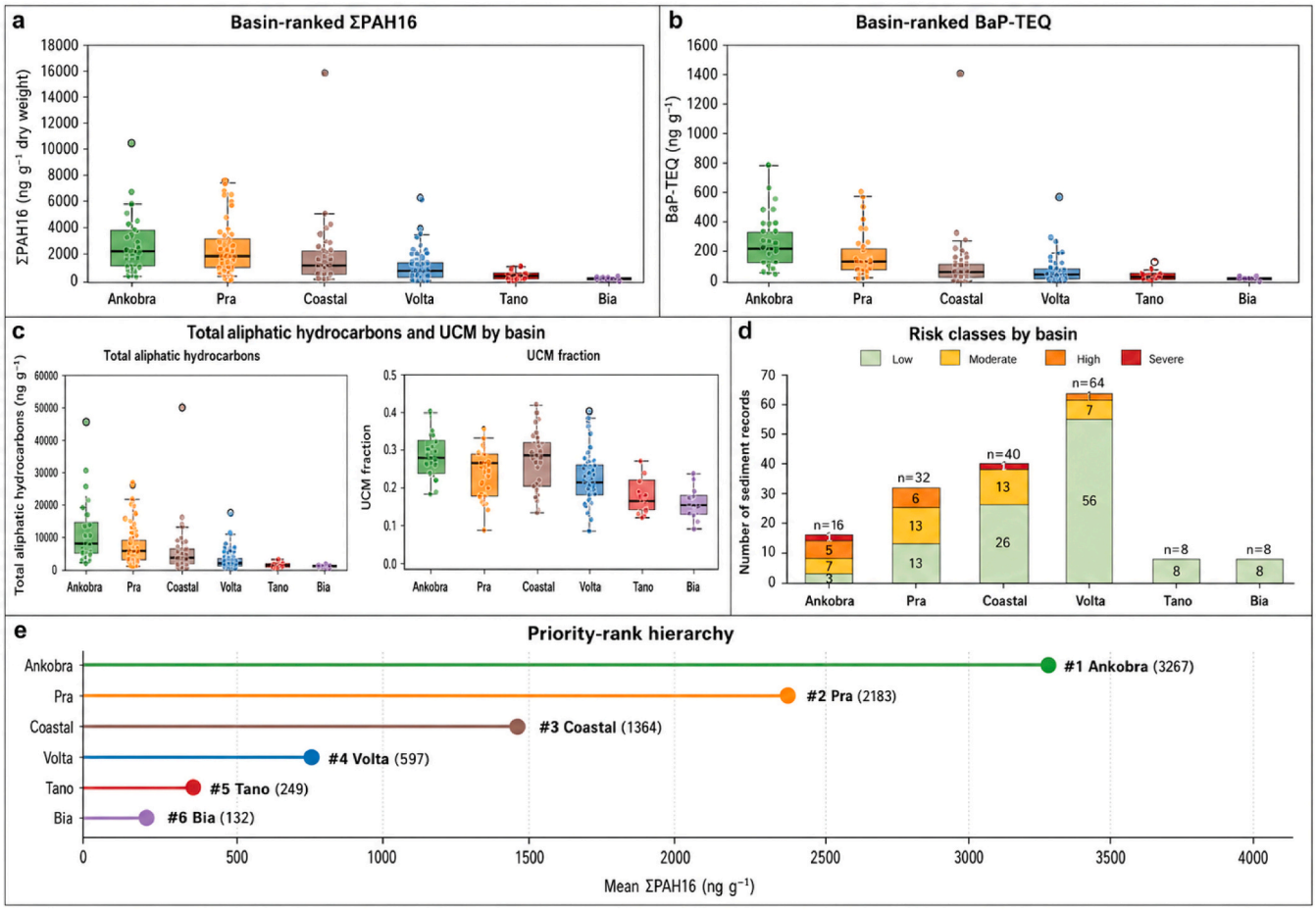



**Fig. 2.** Basin scale hydrocarbon burden and risk hierarchy. Basin level ranking of hydrocarbon contamination burden and ecological risk across Ghanaian river systems. **(a)** Basin ranked distributions of ΣPAH16 concentrations showing highest PAH burdens in Ankobra and Pra, intermediate levels in Coastal systems and lower burdens in Volta, Tano and Bia. **(b)** Basin ranked BaP-TEQ distributions indicating that toxicity weighted PAH burden follows a similar hierarchy, with Ankobra and Pra as dominant toxicity hotspots. **(c)** Total aliphatic hydrocarbon concentrations and UCM fractions by basin, highlighting elevated aliphatic burden and petroleum-weathering signatures in higher pressure basins. **(d)** Stacked risk-class distribution by basin showing the number of low, moderate, high and severe risk sediment records. **(e)** Priority-rank plot summarising the basin hierarchy based on mean ΣPAH16 burden: Ankobra > Pra > Coastal > Volta > Tano > Bia. The figure identifies Ankobra and Pra as the highest-priority hydrocarbon-risk basins, while Tano and Bia represent low burden systems.

**Table 2**
Basin priority ranking for hydrocarbon surveillance.

| Rank | Basin | Rivers represented | n | Mean ΣPAH16 (ng/g) | Mean BaP-TEQ (ng/g) | Mean total aliphatics (ng/g) | High/ severe count | Dominant source signal |
|---|---|---|---|---|---|---|---|---|
| 1 | **Ankobra** | Ankobra, Bonsa | 16 | 3267.3 | 258.8 | 11706.8 | 6 | Mining–urban mixed / pyrogenic |
| 2 | **Pra** | Anum, Birim, Ofin, Pra | 32 | 2183.4 | 191.7 | 6511.0 | 6 | Mining–urban mixed / pyrogenic |
| 3 | **Coastal** | Ayensu, Densu, Kakum, Ochi-Amissah, Tordzie | 40 | 1363.9 | 114.3 | 4883.5 | 1 | Traffic/urban combustion |
| 4 | **Volta** | Afram, Black Volta, Daka, Lower Volta, Oti, Pru, Sene, White Volta | 64 | 597.4 | 51.6 | 1891.6 | 1 | Biomass combustion |
| 5 | **Tano** | Tano | 8 | 249.1 | 21.5 | 802.3 | 0 | Biomass combustion |
| 6 | **Bia** | Bia | 8 | 132.0 | 9.6 | 552.9 | 0 | Terrestrial biogenic / mixed low risk |

Note: This table is a screening priority summary rather than a regulatory ranking. Basin n is unequal, and Coastal and Volta contain high leverage single site maxima despite lower basin means. For this reason, P90 values, maximum concentrations, river level statistics and station level concentrations should be checked in Supplementary Table 7 before management action is taken.

sediment fingerprint.

### 3.5. Latent source apportionment produces a national hydrocarbon-source map

The latent source classification resolved five dominant national source classes: biomass combustion in 59 samples, traffic combustion in 41 samples, mining-urban mixed input in 39 samples, terrestrial biogenic input in 20 samples, and petrogenic input in 9 samples (Fig. 5; Supplementary Table 5). This distribution shows that Ghana's river sediment hydrocarbon burden is controlled by a three-source national triangle: a broad biomass/pyrogenic background, an urban traffic

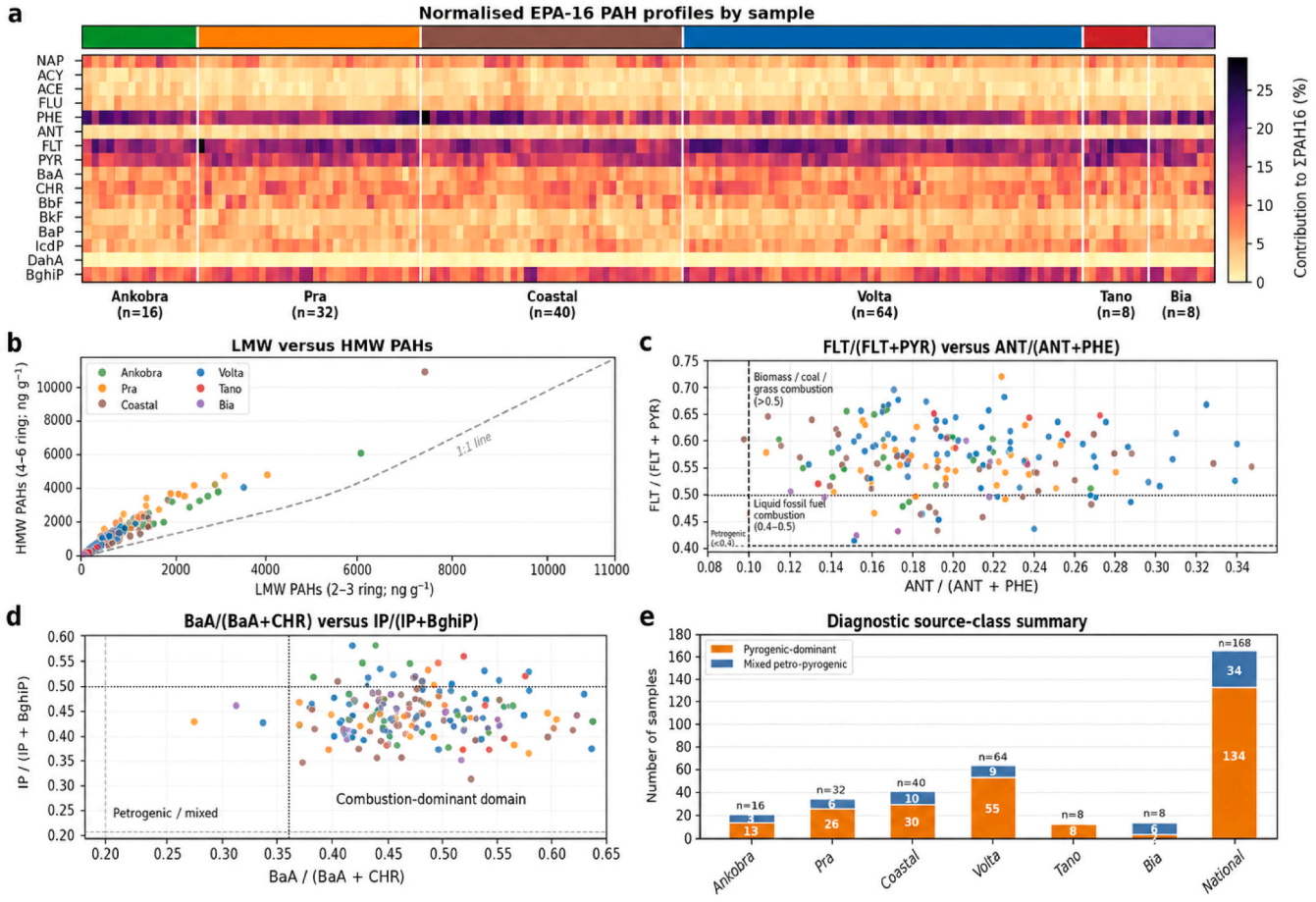



**Fig. 3.** PAH molecular fingerprints and diagnostic ratio forensics. Molecular PAH fingerprints and diagnostic ratio evidence for national source structure. **(a)** Normalised EPA−16 PAH profiles across all sediment samples, clustered by basin system, showing compound level variation in PAH composition. Higher relative contributions of phenanthrene, fluoranthene, pyrene and selected high molecular weight PAHs indicate combustion influenced signatures across much of the dataset. **(b)** LMW versus HMW PAH scatterplot demonstrating widespread enrichment of 4–6 ring PAHs relative to 2–3 ring PAHs, consistent with dominant pyrogenic input. **(c)** Diagnostic cross-plot of FLT/(FLT+PYR) versus ANT/(ANT+PHE), separating petrogenic, liquid fossil-fuel combustion and biomass/coal/grass combustion domains. **(d)** Diagnostic cross-plot of BaA/(BaA+CHR) versus IP/(IP+BghiP), further resolving combustion-dominant and mixed/petrogenic signatures. **(e)** Basin and national summary of diagnostic source classes, showing national pyrogenic dominance with mixed petro-pyrogenic signatures present in selected basins. The figure shows that PAH ratios reveal broad combustion dominance while still detecting mixed petroleum combustion inputs in key hydrocarbon impacted settings.

combustion domain, and a mining urban mixed domain. The spatial organisation of these sources was basin dependent. Biomass combustion dominated much of the Volta and Tano systems, consistent with lower intensity rural and agricultural pressure. Mining-urban mixed sources dominated Ankobra and Pra, matching their higher ΣPAH16, BaP-TEQ, total aliphatic and high/severe risk counts. Traffic combustion was most important in the Coastal basin, especially in Densu and other urban or fuel transport influenced reaches. Terrestrial-biogenic signatures were strongest in Bia and selected low burden reaches, while petrogenic dominance was comparatively rare but important as a forensic indicator of petroleum-related residues. The source risk coupling further indicates that HRSI surveillance risk is not controlled by ΣPAH16 concentration alone. Mining-urban mixed and traffic-combustion classes produced broader and higher HRSI distributions than biomass-combustion and terrestrial-biogenic classes (Fig. 5d). This supports the interpretation that the highest risk national hydrocarbon syndromes are pyrogenic-petroleum-urban mixtures intensified by basin pressure and depositional context.

### 3.6. Reach position and seasonality reveal transport and depositional controls

Hydrocarbon burden increased systematically along the river continuum. Mean ΣPAH16 rose from 609.6 ng $g^{-1}$ in upper/reference reaches to 1073.9 ng $g^{-1}$ in midstream reaches, 1600.3 ng $g^{-1}$ in downstream reaches and 1906.3 ng $g^{-1}$ in estuary/lower depositional reaches (Supplementary Table 7c; Appendix Fig. 6). BaP-TEQ and total aliphatic hydrocarbons followed the same longitudinal trend, and high/severe records increased from 0 in upper/reference reaches to 1 in midstream reaches, 6 in downstream reaches and 7 in lower depositional reaches. This reach gradient is consistent with transport, accumulation and depositional retention influencing hydrocarbon burden. Upper reaches mostly captured background or reference signals, midstream reaches reflected transitional accumulation, downstream reaches captured settlement, mining and market influences, and lower depositional zones functioned as sedimentary traps where PAHs, aliphatic hydrocarbons and UCM-rich residues could accumulate. Because sediment texture and organic-carbon content can also concentrate hydrophobic compounds, the reach effect should be interpreted together with TOC, black carbon and silt-clay fraction. Seasonal contrasts also showed hydrological structuring. Dry-season samples had higher mean ΣPAH16 than wet-season samples (1569.5 versus 1025.5 ng $g^{-1}$), higher mean BaP-TEQ (131.8 versus 87.2 ng $g^{-1}$), higher mean total aliphatics (5295.5 versus 3310.5 ng $g^{-1}$) and more high/severe records (10 versus 4) (Supplementary Table 7d; Appendix Fig. 5). Both severe records occurred during the dry season. The dry-season enrichment suggests

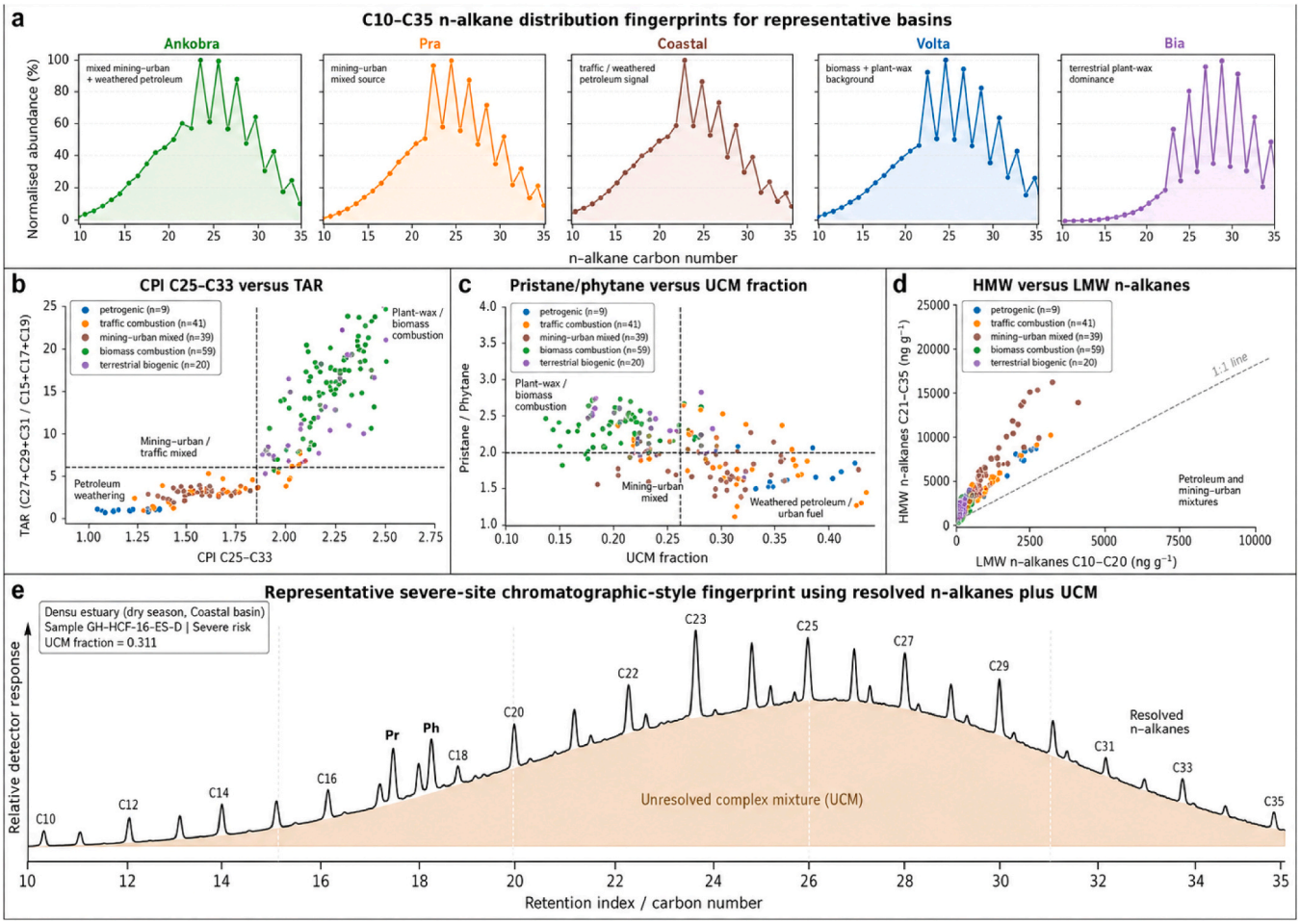



**Fig. 4.** Aliphatic biomarker source disentanglement clarifies source mixtures. Aliphatic hydrocarbon biomarkers used to resolve source mixtures not fully separable from PAH ratios alone. (a) Representative C10–C35 *n*-alkane distribution fingerprints for selected basin systems, showing contrasting hydrocarbon signatures including mining-urban mixtures, weathered petroleum signals, biomass related inputs and terrestrial plant wax dominance. (b) CPI C25–C33 versus TAR scatterplot distinguishing petroleum-weathering, mining-urban/traffic mixed and plant wax/biomass-combustion domains. (c) Pristane/phytane ratio versus UCM fraction showing separation between biomass/plant wax inputs, mining-urban mixtures and weathered petroleum/urban fuel signatures. (d) HMW versus LMW *n*-alkane plot demonstrates widespread high molecular weight *n*-alkane dominance and source dependent separation among petroleum, traffic, mining-urban, biomass and terrestrial-biogenic inputs. (e) Representative chromatographic-style fingerprint from a severe risk site showing resolved *n*-alkane peaks superimposed on a broad unresolved complex mixture, characteristic of weathered petroleum and complex anthropogenic hydrocarbon input. The figure indicates that aliphatic biomarkers add source screening evidence for petroleum weathering, terrestrial plant wax inputs, biomass combustion and mining–urban mixtures that PAH ratios alone cannot fully resolve.

reduced dilution, stronger concentration of sediment-associated hydrocarbons and retention of depositional hotspots during low flow conditions. The wet season reduction in mean burden is consistent with dilution, redistribution and sediment flushing, although basin-level results show that some systems, particularly parts of the Volta, can exhibit local wet season increases due to runoff mobilisation or catchment specific transport effects. The Volta wet season exception is consistent with a first flush and flood scouring interpretation rather than a basin wide reversal of the national dry season pattern. In selected Volta reaches, early wet-season runoff can mobilise road dust, charred biomass residues, agricultural soil particles and bank stored organic matter into depositional zones, while higher discharge can scour and redistribute previously stored contaminated fine sediment. Because the affected Volta records occur within a broad basin that includes agricultural, biomass-burning, transport and reservoir/lower reach settings, the wet season increase is interpreted as a catchment specific mobilisation signal superimposed on a mostly low risk basin background.

### 3.7. Land use pressure covariates identify national source risk predictors

Hydrocarbon burden and HRSI were associated with anthropogenic land-use pressure. Spearman correlations showed positive relationships between ΣPAH16 and mining pressure ($\rho = 0.65$), urban land use ($\rho = 0.56$), road density ($\rho = 0.56$) and population density ($\rho = 0.48$). Similar positive correlations were observed for BaP-TEQ and HRSI. UCM fraction was most strongly associated with urban land use and road density, indicating that petroleum-weathering and unresolved hydrocarbon mixtures are closely linked to urban-transport pressure. In contrast, agriculture and forest cover showed negative correlations with ΣPAH16, BaP-TEQ, UCM fraction and HRSI (Supplementary Table 8; Appendix Fig. 7). Because urban land use, road density and population density were strongly collinear (VIF = 18.05, 9.74 and 6.46), they were treated as a shared urban-traffic pressure complex rather than as independent drivers. Ridge and partial least-squares sensitivity checks retained the same directional interpretation as the transparent OLS screening models. Together, the covariate evidence shows that Ghanaian river-sediment hydrocarbon burden and screening risk are associated with mining pressure, urban-traffic pressure and depositional reach position.

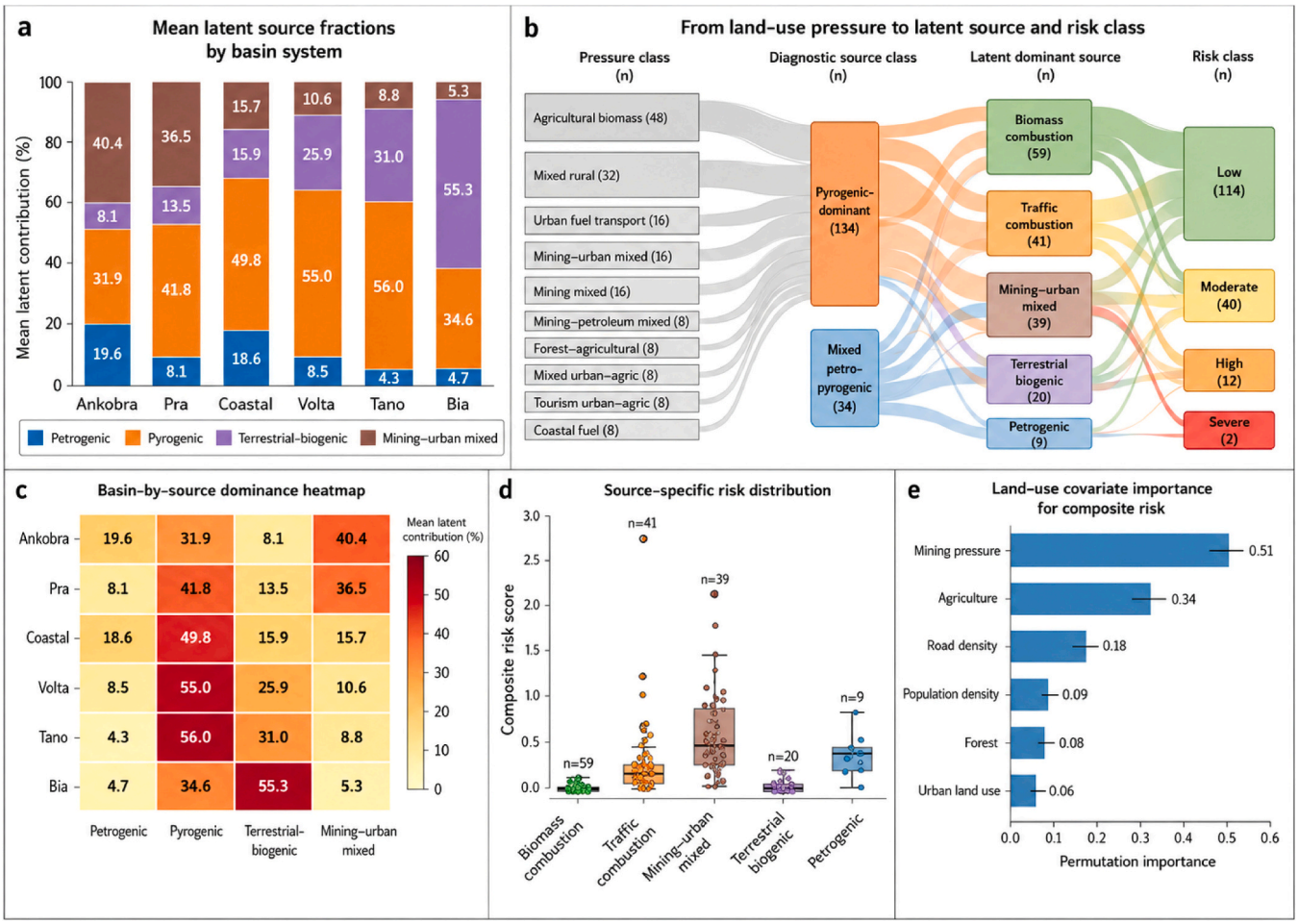



**Fig. 5.** Integrated source apportionment and source risk coupling. Integrated source screening framework linking land use pressure, diagnostic source classification, latent source mixtures and HRSI surveillance risk. (a) Mean latent source fractions by basin system, showing contrasting contributions from petrogenic, pyrogenic, terrestrial-biogenic and mining-urban mixed sources. (b) Sankey/alluvial diagram tracing source pathways from dominant pressure class through PAH diagnostic source class and latent dominant source to final ecological risk class. (c) Basin by source dominance heatmap showing the relative contribution of each latent source component across basin systems. (d) Source specific risk distributions indicating that mining-urban mixed and traffic-combustion sources have broader and higher HRSI distributions than biomass combustion and terrestrial-biogenic sources. (e) Land use covariate importance for HRSI surveillance risk, interpreted cautiously because urban land use, road density and population density are collinear. The figure summarises associations between source mixtures, land use pressure and screening risk rather than proving direct causality.

These are observational associations and should be confirmed with repeated monitoring, source-material fingerprints and toxicity testing. Partial correlation analysis supported a dual source input and depositional enrichment interpretation. The positive association between ΣPAH16 and mining or urban-traffic pressure persisted after controlling for TOC and silt-clay fraction, although the effect was attenuated. Conversely, TOC and silt-clay fraction remained positively associated with ΣPAH16 and BaP-TEQ after controlling for anthropogenic pressure. Thus, high concentrations reflect both stronger source input and greater retention on organic rich, fine grained sediment, which explains why high risk classes are concentrated in depositional or transitional reaches.

## 4. Discussion

### *4.1. Ghanaian rivers do not share a single hydrocarbon problem*

The central implication of this national dataset is that Ghanaian river sediments do not express one uniform hydrocarbon contamination problem. Instead, the hydrocarbon atlas resolves several basin-specific regimes, differentiated by PAH burden, BaP-equivalent toxicity, aliphatic hydrocarbon structure, UCM enrichment, latent-source dominance and screening risk class. This is important because a single national average would obscure the fact that Ankobra, Pra, Coastal, Volta, Tano and Bia systems occupy different positions along a source risk continuum. The basin-priority ranking places Ankobra first, followed by Pra, Coastal systems, Volta, Tano and Bia, but this ordering should be read with the sample-size imbalance, hotspot influence and screening level risk design in mind. The first regime is a mining-urban high risk regime, expressed most strongly in the Ankobra and Pra systems and in rivers such as Ankobra, Bonsa, Birim, Ofin and the Pra mainstem. This regime is characterised by high ΣPAH16, elevated BaP-TEQ, high total aliphatic hydrocarbons, strong UCM contribution and a dominant mining-urban latent source. The Ankobra basin recorded a mean ΣPAH16 of 3267.3 ng $g^{-1}$ , mean BaP-TEQ of 258.8 ng $g^{-1}$ , mean total aliphatics of 11,706.8 ng $g^{-1}$ and six high/severe records, while Pra recorded a mean ΣPAH16 of 2183.4 ng $g^{-1}$ and six high/severe records. At river scale, Ankobra, Bonsa, Birim, Ofin and Pra carry the clearest mining-urban signatures, confirming that hydrocarbon risk is concentrated in river corridors where extractive activity, settlement, transport and fuel use overlap. The second regime is a traffic/urban combustion regime, represented principally by the Densu, Lower Volta and selected Coastal rivers. This regime is dominated by high-molecular-weight PAHs, traffic-combustion latent fractions, urban land-use pressure and petroleum-weathering indicators. The Densu severe hotspot is especially

important because it demonstrates that coastal systems may have lower basin-wide mean values than Ankobra and Pra yet still generate the most extreme single-site hazard where urban traffic, fuel handling and estuarine depositional trapping coincide. Appendix Fig. 8 explicitly identifies the Densu severe site as an estuarine traffic-combustion hotspot with extreme PAH toxicity and aliphatic burden, while the Lower Volta appears as an isolated petroleum/traffic signal within an otherwise lower-priority basin. The third regime is a biomass-combustion background regime, expressed mainly across Volta tributaries and agricultural or mixed-rural basins. This regime is still pyrogenic in PAH-ratio space, but its ecological significance is lower because ΣPAH16, BaP-TEQ, total aliphatic burden and HRSI values are generally modest. Volta therefore should not be interpreted as uncontaminated; rather, it represents a broad national background system with mostly low risk and isolated hotspots. Its large spatial coverage and heterogeneous pressure structure mean that surveillance should be maintained even where the dominant source appears to be biomass combustion. The fourth regime is a terrestrial-biogenic low risk regime, best expressed in Bia and selected forest/agricultural systems. This regime has low ΣPAH16, low BaP-TEQ, low total aliphatic hydrocarbons and strong terrestrial-biogenic or plant wax signatures. Bia recorded the lowest basin mean ΣPAH16, lowest BaP-TEQ and no high/severe records, making it a useful low-burden reference or early warning basin rather than a priority remediation basin. Tano also occupies the low-burden end of the national gradient, although its dominant signal is closer to biomass combustion than the stronger terrestrial-biogenic signal observed in Bia.

### 4.2. Comparison with African and global sediment PAH

Comparison with available African and global sediment studies places Ghanaian river sediments in a broad low-to-hotspot contamination spectrum rather than a single national contamination class. A compact benchmarking matrix has been added as Supplementary Table 11 to make this comparison explicit. Within the Ghanaian dataset, low-burden Bia and Tano basin means (132.0 and 249.1 ng $g^{-1}$, respectively) are closer to low or moderate coastal and lagoonal sediment contexts, whereas Ankobra, Pra and the Densu/Ankobra severe records are closer to urbanised estuarine, mangrove, mining-influenced or petroleum-impacted hotspot settings reported in West Africa and Asia [2–4,29]. The severe Ghanaian records (ΣPAH16 = 10,378.74–15, 895.6 ng $g^{-1}$) occupy the upper hotspot range of the comparative literature, while most Ghanaian samples remain low or moderate. The comparison is intentionally conservative because analyte lists, sediment fractions, extraction procedures, reporting units, hydrological seasons and risk tools differ among studies. Its purpose is therefore to establish order-of-magnitude context and surveillance relevance, not to imply direct regulatory equivalence across unrelated sediment datasets. This contrast reinforces the risk-compression interpretation: the national issue is not uniformly high PAH loading, but a small number of river reaches where anthropogenic input, petroleum-weathering signal and sediment capture converge.

### 4.3. PAH monitoring is insufficient for Ghanaian river systems

A major conceptual finding is that PAH-only monitoring would oversimplify Ghana's hydrocarbon problem. The PAH diagnostic ratios classify most samples as pyrogenic-dominant, which is scientifically meaningful because fluoranthene, pyrene and other high molecular weight PAHs are widespread across the dataset. However, this classification compresses multiple source processes into a single broad combustion label. Diagnostic-ratio plots separate petrogenic, mixed and pyrogenic domains, but the appendix also shows diagnostic uncertainty and threshold proximity effects, meaning that ratio based assignments should be interpreted cautiously in mixed source tropical catchments. This limitation is consistent with the broader PAH forensics literature. PAH ratios are useful screening tools, but they may be altered by weathering, photodegradation, transport, sediment sorting, combustion-temperature variability and source mixing [13,14]. In the Ghanaian dataset, this limitation is not theoretical: PAH ratios identify widespread pyrogenic dominance, whereas aliphatic biomarkers reveal additional information on petroleum weathering, plant wax dilution, unresolved hydrocarbon mixtures and mining-urban inputs. Thus, a PAH only assessment would correctly identify combustion influence but under detect the petroleum-weathering and mixed-source structure that controls several high-risk sites. The added value of the aliphatic domain is therefore central to the paper's novelty. CPI C25–C33 and TAR separate terrestrial plant-wax and biomass-related signatures from petroleum-weathered or mining-urban mixtures; pristane/phytane and UCM fraction differentiate redox, biodegradation and weathered petroleum signals; and C10–C35 n-alkane distributions expose chain-length fingerprints that are invisible in PAH-only analysis. Fig. 4 explicitly shows that aliphatic biomarkers distinguish mining-urban mixtures, weathered petroleum, traffic/urban-fuel signatures, biomass inputs and terrestrial plant wax dominance, while Fig. 5 shows that latent source apportionment links these molecular fingerprints to basin pressure and ecological risk. This supports the central framing of the manuscript as PAH-aliphatic biomarker source screening, rather than conventional PAH monitoring alone. It also aligns with recent Ghanaian offshore work showing that integrated aliphatic-PAH fingerprinting, Bayesian source apportionment and eco-probabilistic risk assessment can improve source discrimination relative to PAH measurement alone [7]. The present study extends that logic inland, from offshore petroleum influenced sediments to a national river basin system, where mining, biomass burning, traffic, urban runoff and depositional trapping interact within the same sediment archive.

### 4.4. National HRSI mapping is more informative than isolated threshold comparison

The HRSI results were interpreted as a comparative national surveillance framework. HRSI was explicitly designed to rank sites and basins for surveillance priority, not to function as a statutory threshold, toxicity standard or site-specific ecological effect prediction. Its strength is that it combines three complementary dimensions of hydrocarbon hazard: ΣPAH16 burden, BaP-equivalent toxicity and UCM/petroleum-weathering burden. This approach is more informative than isolated threshold comparison because Ghana's hydrocarbon risk is spatially compressed. Most samples fall in the low or moderate classes, but the national hazard is disproportionately controlled by a small number of intense hotspots. The two severe samples Densu estuary in the dry season and Ankobra midstream in the dry season represent different source risk syndromes: one traffic-combustion and estuarine-depositional, the other mining-urban mixed and pyrogenic. Therefore, the key management problem is not uniform national contamination, but the identification of basin-reach-season combinations where source intensity and depositional retention amplify ecological risk. The reach and seasonal patterns are consistent with this interpretation. Mean ΣPAH16 increases from upper/reference reaches to downstream and estuary/lower depositional reaches, while high/severe records are concentrated in downstream and lower depositional settings. Dry season samples also show higher national mean ΣPAH16, BaP-TEQ, total aliphatics and high/severe counts than wet-season samples, consistent with low flow concentration, reduced dilution and depositional retention. These interpretations remain partly inferential because hydrodynamic measurements, sedimentation rates and repeated multi-year sampling were not included. Mechanistically, risk compression reflects the coupling of source pressure with sedimentary retention. Mining and urban–traffic inputs supply PAHs and petroleum-weathered residues, while low gradient downstream, estuarine and reservoir/lower depositional reaches retain organic rich fine sediment and UCM bearing particles. The same process explains why the two severe records occur in dry season depositional or transitional settings and why selected Volta reaches can

show wet season increases when runoff and flood scouring mobilise stored catchment material. The term sediment capture is therefore used to describe this combined hydrological, geomorphological and geochemical retention process. The national risk map therefore performs a different function from a conventional concentration table. It organises molecular evidence into a basin-prioritised surveillance tool by distinguishing toxicity weighted PAH hotspots, petroleum-weathering/UCM hotspots and source class hotspots. Fig. 6 is policy facing because it integrates hydrocarbon burden, toxicity, source certainty and high/severe counts into a basin priority framework for targeted monitoring and adaptive environmental management.

### 4.5. Land use pressure is associated with hotspot occurrence

The national source risk patterns are associated with land-use pressure rather than proven by land use alone. Mining pressure and the urban–traffic pressure index were positively associated with ΣPAH16, BaP-TEQ, UCM fraction and HRSI, while agriculture and forest cover showed negative associations. This means that the high risk hydrocarbon syndromes are spatially aligned with catchment pressure, fuel use, transport intensity, settlement density, extractive activity and sediment-retention environments. The wording is intentionally association-based because the study design was observational and did not directly measure emission rates or transport fluxes. The regression results further indicate that mining pressure retained the clearest positive association with ΣPAH16 and BaP-TEQ in the multiple models. Urban land use, road density and population density were not interpreted separately because their high VIF values indicated strong multicollinearity; instead, the urban–traffic pressure index provided a more defensible summary of their shared signal. Ridge regression and partial least-squares sensitivity check reduced coefficient instability while preserving the same broad associations. Agriculture and forest cover were associated with lower hydrocarbon burden, likely because their dominant sedimentary signal is diluted by biomass derived or terrestrial-biogenic organic matter rather than concentrated petroleum-combustion inputs. These relationships provide plausible support for the four regime interpretation, but they do not prove a single causal pathway. Ankobra and Pra are high priority because mining pressure, settlement activity, road linked fuel use and downstream depositional accumulation overlap. Densu type coastal hotspots are consistent with severe HRSI surveillance risk where urban traffic emissions, fuel residues and lower depositional zones coincide. Volta is mostly low risk because much of its spatial coverage reflects biomass-combustion or mixed-rural backgrounds, although isolated urban or petroleum/traffic nodes still require surveillance. Bia and Tano remain low burden because their dominant signatures are terrestrial-biogenic or biomass related and because high intensity mining-urban or traffic-fuel inputs are less strongly expressed in the sampled reaches.

### 4.6. Study limitations and uncertainty

First, the study used one dry season and one wet-season sampling campaign, so interannual variability and event driven first flush pulses were not captured. Second, the analysis was based on surface sediment

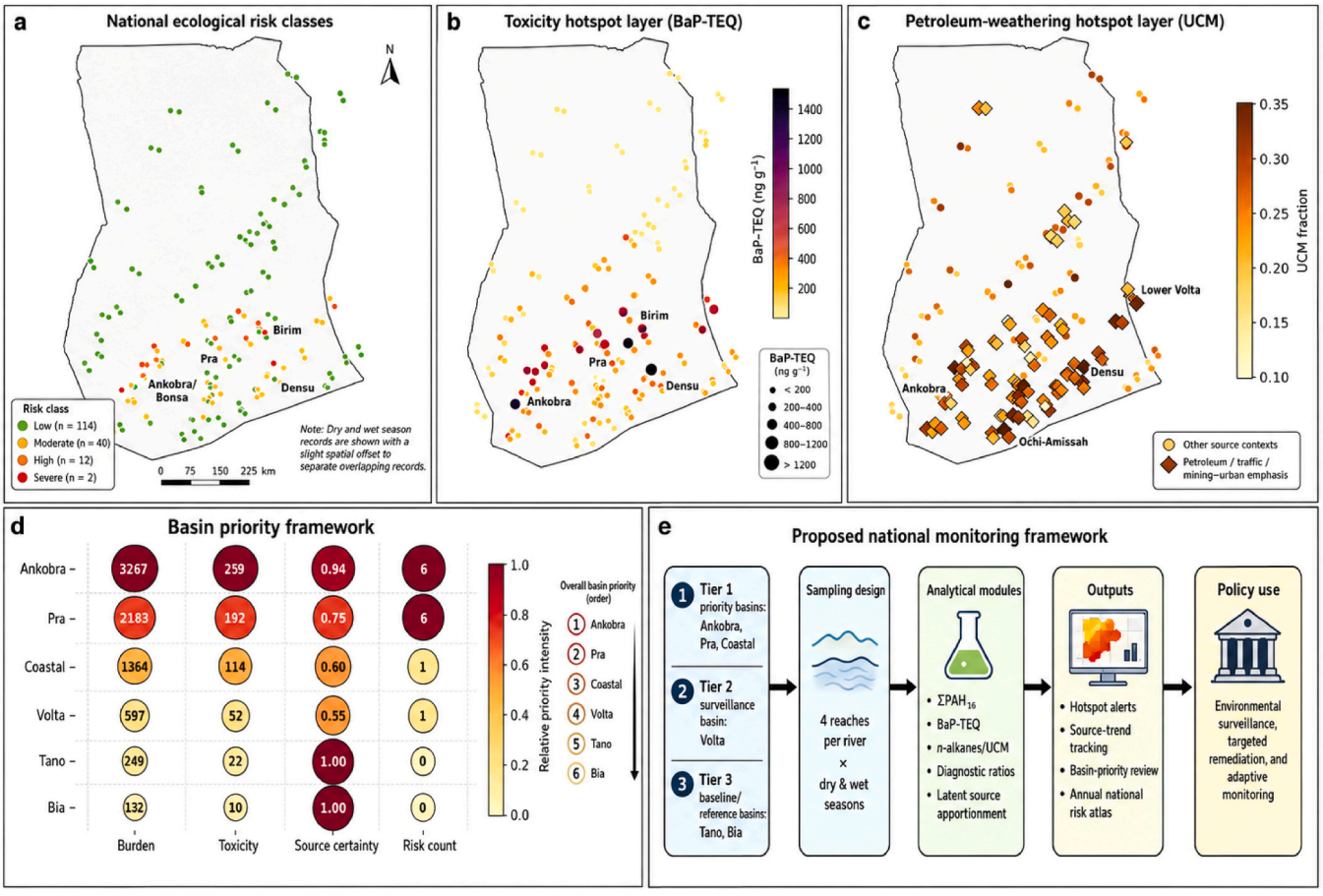



**Fig. 6.** National ecological risk map and basin priority framework. Screening level national hydrocarbon risk atlas based on the Hydrocarbon Risk Surveillance Index (HRSI). (a) National map of HRSI classes across Ghanaian river sediment sampling locations. (b) Basin priority framework integrating ΣPAH16, BaP-TEQ, UCM/petroleum-weathering burden and high/severe HRSI counts. (c) Sensitivity checks showing that priority ranking is robust to increased BaP-TEQ weighting and to high leverage hotspot evaluation.

only; it did not include sediment cores, sedimentation rates, continuous discharge records or direct hydrodynamic measurements. Third, HRSI classes are surveillance categories based on chemical burden, BaP-TEQ and UCM structure, not biological toxicity tests or statutory regulatory thresholds. Fourth, latent source classes are model dependent and should be verified with local source materials such as fuel residues, road dust, urban drain sediments, mining-area sediments and biomass-burning residues. Fifth, urban land use, road density and population density were highly collinear, so their individual effects cannot be cleanly separated; the urban-traffic pressure index is a conservative compromise rather than proof of mechanism. Sixth, partial correlations suggest both source input and depositional enrichment, but causal separation requires catchment scale flux data, controlled source fingerprints and repeated hydrological sampling. Finally, the Densu and Ankobra severe records are high leverage observations; they are real surveillance priorities, but basin scale conclusions should be checked through repeat sampling and targeted local investigation.

### *4.7. Policy and monitoring implications for Ghana*

The first monitoring implication is that Ankobra and Pra require priority surveillance and confirmatory monitoring. These basins have the highest mean ΣPAH16, highest BaP-TEQ, elevated total aliphatic burdens, strong UCM signal and the largest high/severe HRSI counts. Monitoring should prioritise Ankobra, Bonsa, Birim, Ofin and the Pra mainstem, with special attention to downstream and estuary/lower depositional reaches where risk is amplified by sediment capture. Second, Densu type coastal urban hotspots require event based and dry season monitoring. The Densu severe record demonstrates that the Coastal operational system can produce extreme local risk even when basin wide averages are lower than Ankobra and Pra. For urban coastal rivers, routine seasonal sampling should therefore be supplemented with targeted monitoring after fuel spills, road-runoff events, market or transport intensification, storm-drainage discharge and low flow concentration periods. Third, Volta requires broad surveillance rather than blanket low priority classification. Although Volta is mostly low risk and biomass-combustion dominated, its large spatial coverage, wet season mobilisation potential and isolated petroleum/traffic anomalies require distributed monitoring. Fourth, Bia and Tano can function as low-burden reference or early-warning systems. Finally, future Ghanaian sediment monitoring should measure PAHs and aliphatic biomarkers together and report HRSI, ERL/ERM or TEL/PEL comparisons, QA/QC metadata, TOC, particle size and local pressure covariates in one reproducible surveillance workflow.

## 5. Conclusion

This study presents a national PAH-aliphatic hydrocarbon source disentanglement and HRSI framework for Ghanaian river sediments. By integrating PAH distributions, C10–C35 *n*-alkanes, pristane/phytane, CPI, TAR, UCM structure, latent source classification and Hydrocarbon Risk Surveillance Index screening, the work shows that Ghana's river systems do not share a single hydrocarbon contamination pattern. Instead, they form distinct basin specific source risk regimes: Ankobra and Pra are dominated by mining-urban and pyrogenic high burden signatures; Coastal rivers contain intense traffic/urban-combustion hotspots; Volta represents a broad biomass-combustion background with isolated anomalies; Tano is a low burden biomass-influenced system; and Bia functions as a low risk terrestrial-biogenic endmember. PAH diagnostic ratios correctly identify broad combustion dominance, but they compress diverse sources into general pyrogenic classes. The addition of aliphatic biomarkers reveals petroleum weathering, plant wax inputs, mining-urban mixtures and traffic related fuel residues that are otherwise under resolved. The national risk structure is also uneven. Most sediment records indicate low to moderate HRSI concern, while a small number of dry season depositional or transitional reaches carry disproportionate hydrocarbon burden and toxicity equivalent load. The Densu estuary and Ankobra midstream severe-risk sites show where urban traffic, fuel handling, mining pressure and sediment capture may combine to produce localised hazard within broader basin-scale gradients. The results support tiered national surveillance: Ankobra, Pra and selected Coastal hotspots should receive priority monitoring; Volta should remain under broad spatial surveillance because of isolated petroleum/traffic anomalies and wet-season mobilisation potential; and Tano and Bia can serve as low burden reference and early-warning basins. The framework is transferable only as a testable template. Application to other tropical river networks should recalibrate source profiles, sediment properties, hydrology, 1 km or catchment scale land use metrics and HRSI thresholds, and should include repeat sampling, source material confirmation and biological toxicity evidence.

## CRediT authorship contribution statement

**Ibrahim Kwame Kwarteng:** Writing – review & editing, Validation, Resources, Investigation. **Archibold Buah-Kwofie:** Writing – review & editing, Visualization, Validation, Software, Methodology. **Zainab Ngaini:** Writing – review & editing, Supervision, Project administration, Methodology. **Rafeah Wahi:** Writing – review & editing, Supervision, Formal analysis. **Elsie Effah Kaufmann:** Writing – review & editing, Supervision, Project administration, Investigation. **Dickson Abdul-Wahab:** Writing – review & editing, Visualization, Methodology, Investigation. **Ebenezer Aquisman Asare:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Joseph Richmond Fianko:** Writing – review & editing, Software, Investigation. **Crentsil Kofi Bempah:** Writing – review & editing, Software, Methodology.

## Author contributions

EAA conceived of the study, the investigation and drafting of the study. EAA, DA-W, EEK, RW, ZN, ABK, IKK, CKB and JRF edited the study. The author(s) read and approved the final study.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

The authors gratefully acknowledge the Nuclear Chemistry and Environmental Research Centre, National Nuclear Research Institute, Ghana Atomic Energy Commission, for institutional support and scientific input relevant to environmental interpretation and modelling.

## Appendix




**Fig. 1.** Full sampling metadata and QA/QC summary

Summary of the national sediment-sampling metadata and analytical QA/QC architecture used for the hydrocarbon source-disentanglement dataset. The figure summarises the distribution of sediment records across basin systems, river systems, reach positions and seasons, and reports field duplicates, equipment blanks, procedural blanks, method blanks, calibration checks, detection limits, recovery, duplicate RPD and QA/QC grades.

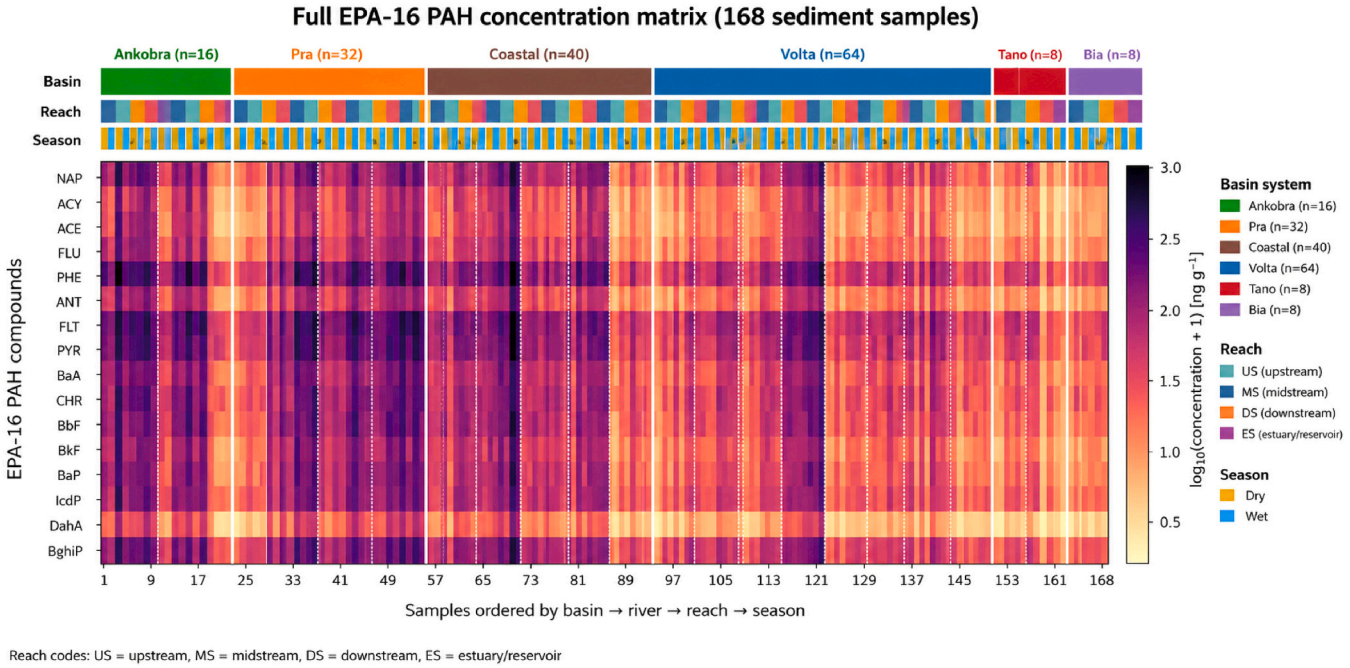



**Fig. 2.** Full EPA−16 PAH concentration matrix

Heatmap of EPA−16 PAH concentrations across all 168 surface sediment samples. Samples are ordered by basin system, river, reach position and season, with annotation strips showing basin, reach and seasonal grouping. Colour intensity represents log-transformed PAH concentration, allowing compound level comparison across the full national dataset. The matrix shows strong spatial heterogeneity in PAH burden, with higher intensity profiles concentrated in Ankobra, Pra and selected Coastal and Volta hotspot samples, while Tano and Bia show consistently lower PAH concentrations.

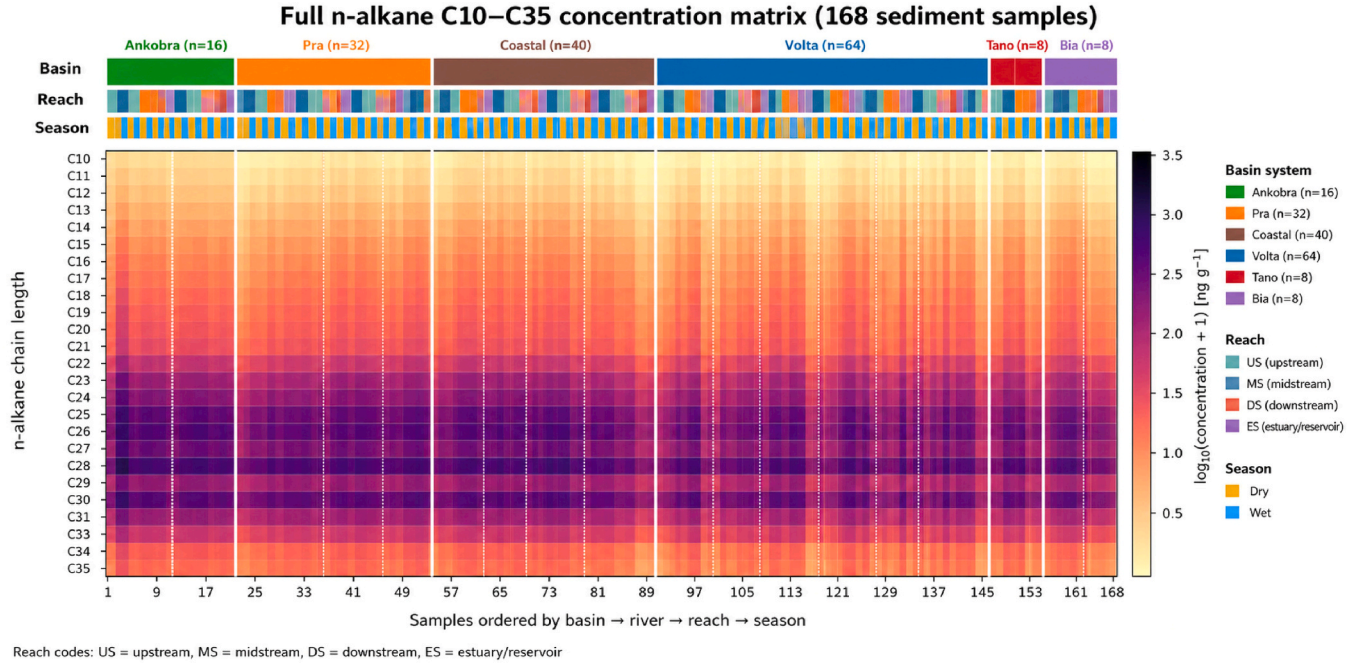



**Fig. 3.** Full *n*-alkane C10–C35 concentration matrix

Heatmap of C10–C35 *n*-alkane concentrations across all 168 sediment records. Samples are arranged by basin system, river, reach position and season to show chain length distributions along the national river continuum. The matrix highlights source dependent aliphatic hydrocarbon fingerprints, including elevated mid to high carbon number *n*-alkanes in higher pressure basins and stronger plant wax/terrestrial signals in lower burden systems. The figure supports the use of aliphatic biomarkers as an additional forensic layer beyond PAH only interpretation.

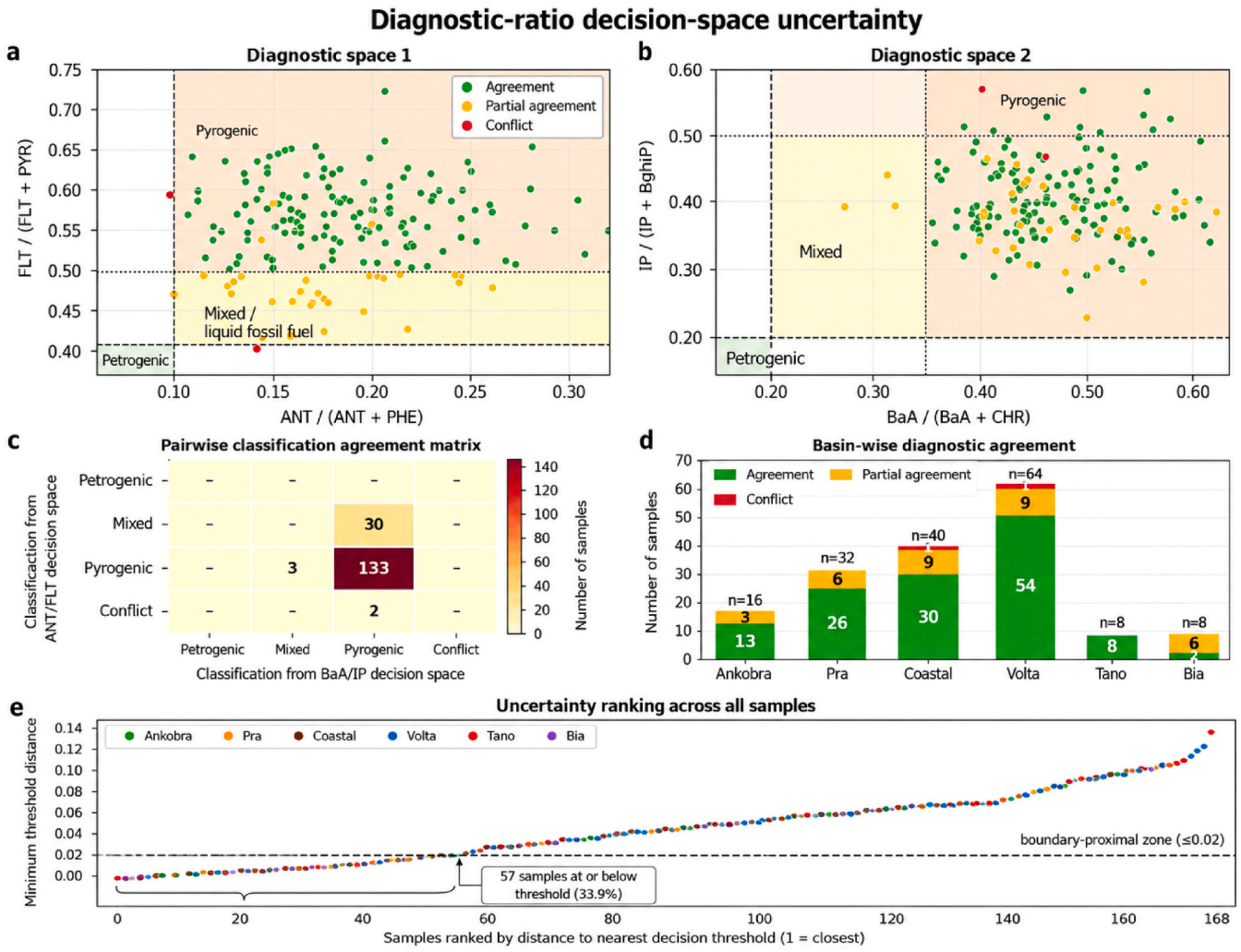



**Fig. 4.** Diagnostic ratio decision space uncertainty

Decision space analysis showing agreement and uncertainty among PAH diagnostic ratio classifications. Cross plots of ANT/(ANT+PHE), FLT/(FLT+PYR), BaA/(BaA+CHR) and IP/(IP+BghiP) separate petrogenic, mixed and pyrogenic domains. The pairwise agreement matrix summarises where diagnostic ratio classifications are consistent or conflicting, while basin level agreement bars show the distribution of agreement, partial agreement and conflict across the six basin systems. The uncertainty ranking identifies samples located close to diagnostic thresholds, indicating where source classification should be interpreted cautiously.

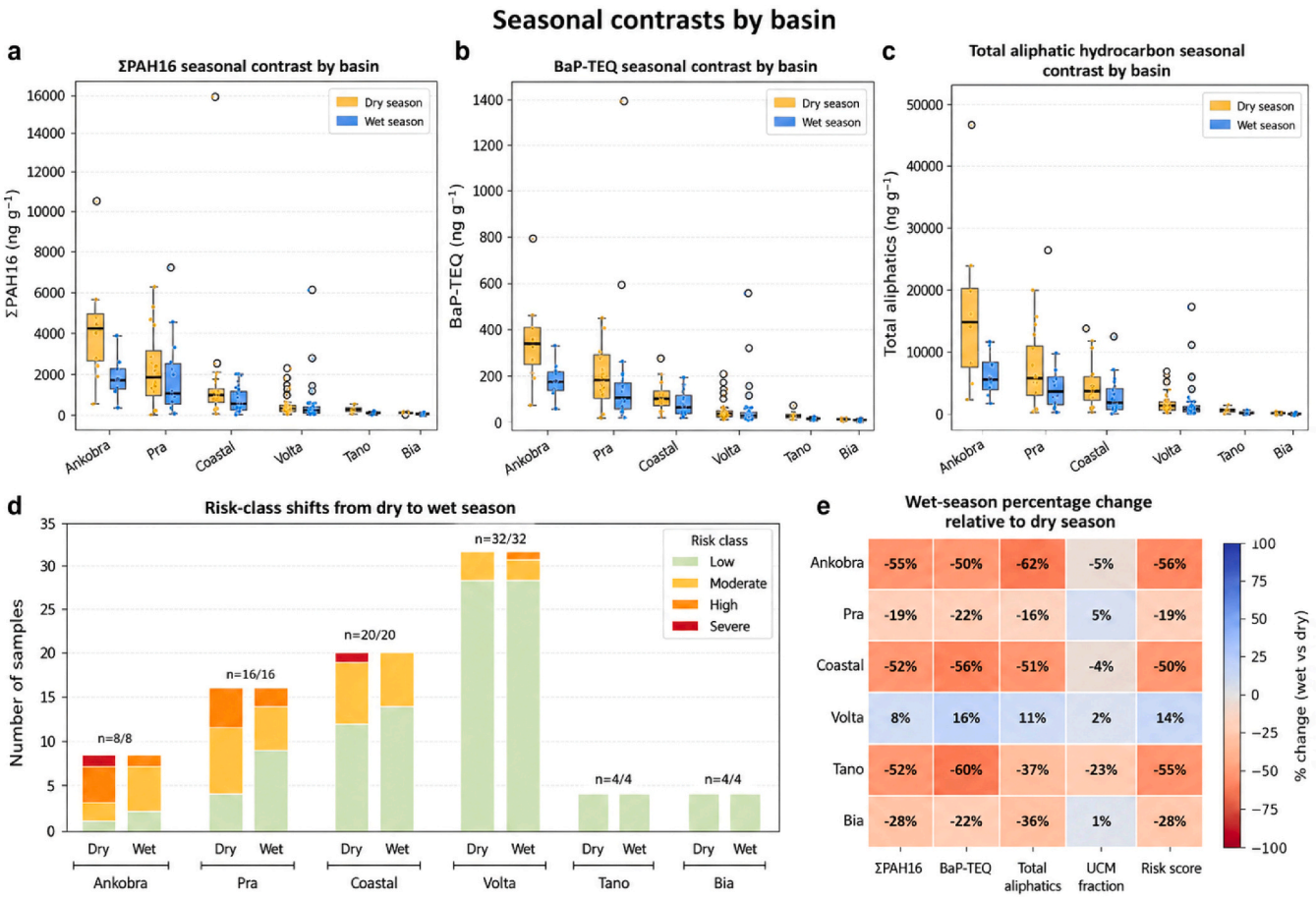



**Fig. 5.** Seasonal contrasts by basin

Dry versus wet season contrasts in hydrocarbon concentration and ecological risk by basin system. Basin specific boxplots compare seasonal distributions of ΣPAH16, BaP-TEQ and total aliphatic hydrocarbons, while stacked bars show dry to wet shifts in risk class composition. The percentage change heatmap summarises seasonal differences in ΣPAH16, BaP-TEQ, total aliphatics, UCM fraction and HRSI. The figure shows that several high pressure basins, especially Ankobra and Coastal systems, exhibit lower wet season mean concentrations consistent with dilution or redistribution, whereas Volta shows modest wet season increases in selected metrics.

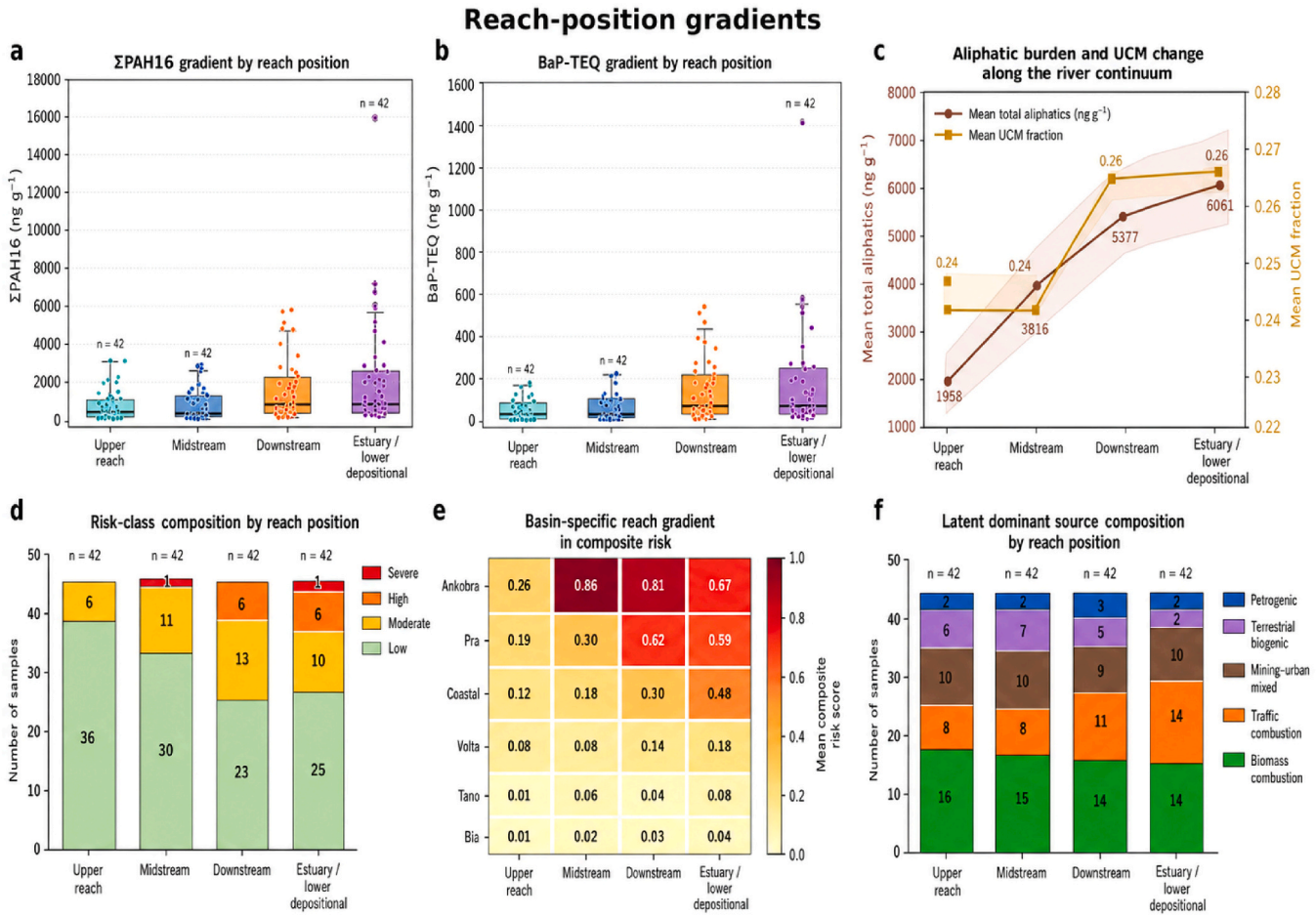



**Fig. 6.** Reach position gradients

Longitudinal contrasts in hydrocarbon burden, toxicity, source indicators and risk from upper/reference reaches through midstream, downstream and estuary/lower depositional zones. Boxplots show reach position gradients in ΣPAH16 and BaP-TEQ, while paired line plots summarise changes in total aliphatic hydrocarbons and UCM fraction along the river continuum. Stacked bars show risk class composition and latent source dominance by reach position, and the basin by reach heatmap identifies where downstream or depositional reaches carry the strongest HRSI risk. The figure indicates that hydrocarbon burden and HRSI surveillance risk generally increase from upper reaches toward downstream and lower depositional zones, with basin specific exceptions.

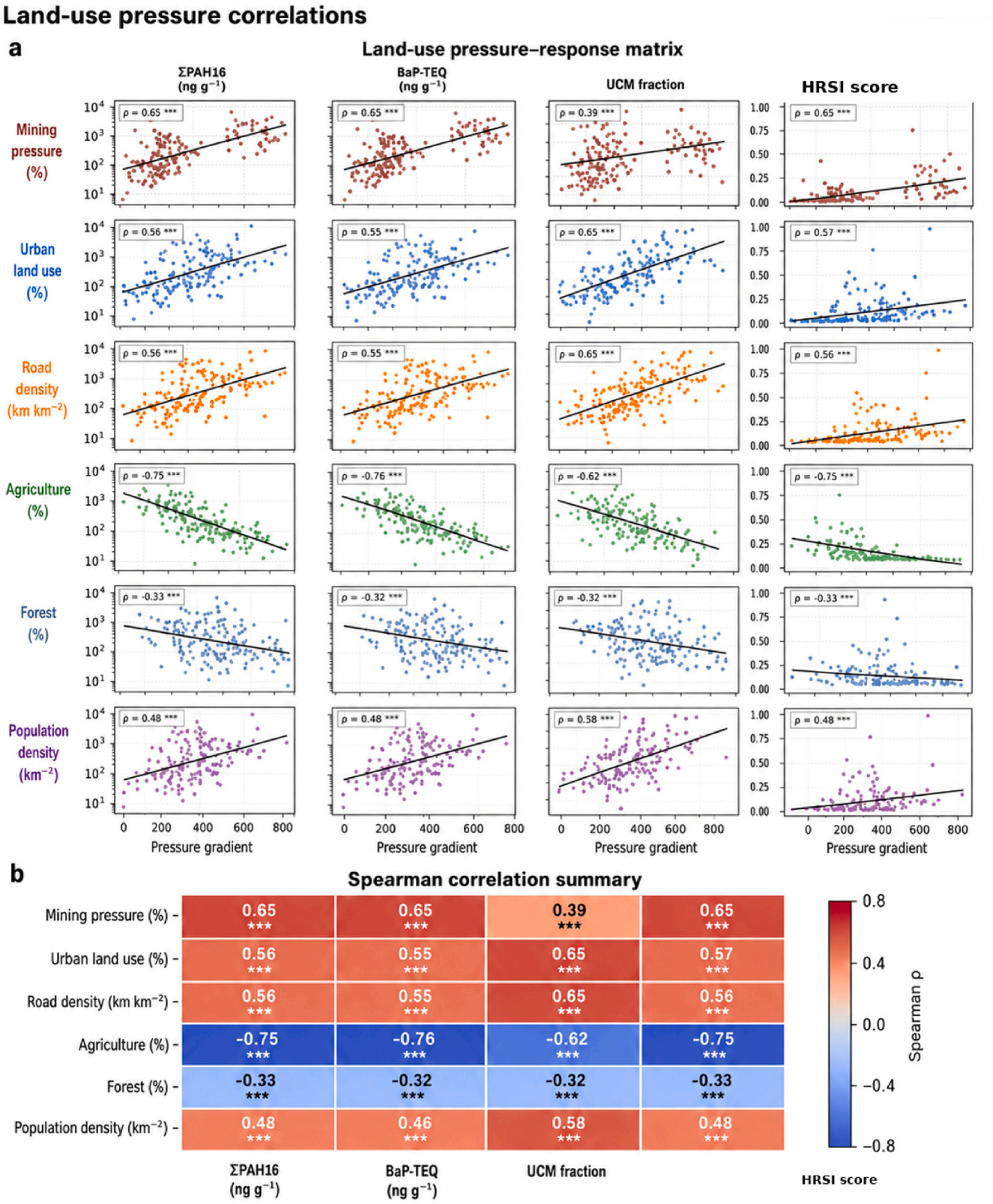



**Fig. 7.** Land use pressure correlations

Relationships between land use pressure variables and hydrocarbon response indicators across the national sediment dataset. Scatterplot matrices show associations between mining pressure, urban land use, road density, population density, agriculture, forest cover and hydrocarbon responses.

The analysis reports the urban–traffic pressure index, VIF diagnostics, ridge/partial least squares sensitivity checks and partial correlations among TOC, silt-clay fraction and ΣPAH16 to separate source input and depositional enrichment signals.

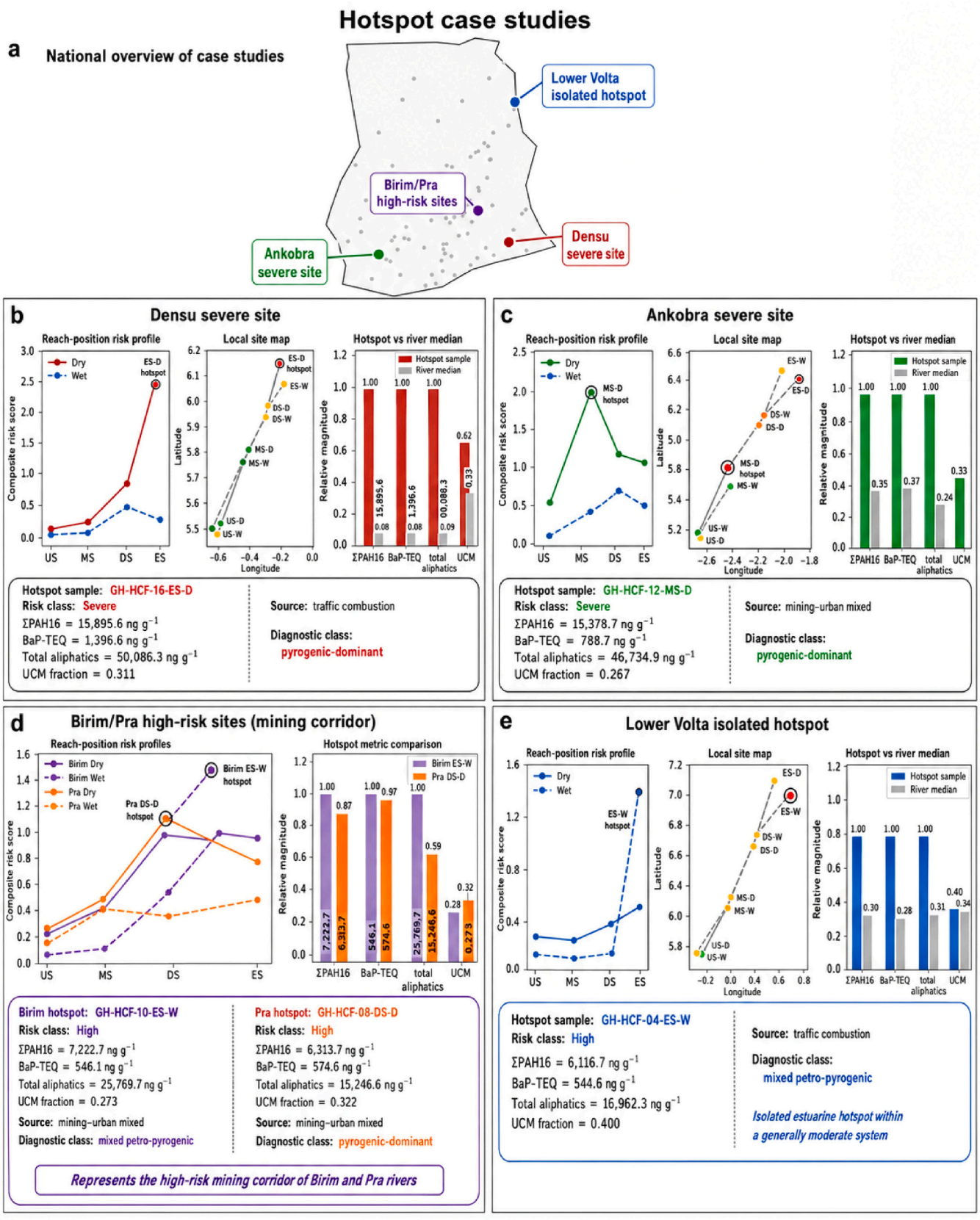



**Fig. 8.** Hotspot case studies

Detailed case study analysis of selected national hydrocarbon hotspots: the Densu severe site, Ankobra severe site, Birim/Pra high risk

mining corridor and Lower Volta isolated hotspot. The interpretation links each hotspot to HRSI class, source disentanglement evidence, land use pressure, reach position, sediment capture and seasonal hydrological behaviour, including the Volta wet-season first flush/flood scouring exception.

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.hazmo.2026.100034.

## Data Availability

Data will be made available on request.

## References


[1] E. Zafeiraki, E. Moulas, K.M. Kasiotis, E. Bakeas, E. Dassenakis, Polycyclic aromatic hydrocarbons (PAHs) in surface sediments from Greece: occurrence, sources and risk assessment, Mar. Pollut. Bull. 197 (2023) 115715.

[2] O.K. Berezi, S.S. Aghomi, O.J. Eriegha, Characterization of polycyclic aromatic hydrocarbons in mangrove sediments from Ifiekporo Creek, Warri, Nigeria, Discov. Environ. 2 (2024) 25.

[3] F. Zhang, D. Zhang, H. Lou, X. Li, H. Fu, X. Sun, Y. Wang, J. Liu, Distribution, sources and ecological risks of PAHs and n-alkanes in water and sediments of typically polluted estuaries: insights from the Xiaoqing River, J. Environ. Manag. 364 (2024) 121471.

[4] D. Cui, Y. Xu, Z. Hou, H. Xu, PAHs in subtropical estuarine sediments of western Beibu Gulf, China: distribution, source apportionment, and ecological risk assessment, Front. Mar. Sci. 12 (2025) 1700618.

[5] E. Gilbert, D.K. Dodoo, F. Okai-Sam, D.K. Essumang, E.K. Quagraine, Characterization and source assessment of heavy metals and PAHs in sediments of the Fosu Lagoon, Ghana, J. Environ. Sci. Health Part. A 41 (12) (2006) 2747–2775.

[6] J. Amoako, O.D. Ansa-Asare, A.Y. Karikari, G. Dartey, Levels of polycyclic aromatic hydrocarbons in the Densu River Basin of Ghana, Environ. Monit. Assess. 174 (2011) 471–480.

[7] E.A. Asare, D. Abdul-Wahab, A. Asamoah, S.B. Dampare, E.E. Kaufmann, R. Wahi, Z. Ngaini, C.K. Klutse, C.K. Bempah, J.R. Fianko, G.A. Gyasi, Bayesian source apportionment and eco-probabilistic risk assessment of aliphatic and polycyclic aromatic hydrocarbons in sediments from the offshore area between the Deep Water Tano and West Cape Three Points blocks, Ghana, Mar. Pollut. Bull. 210 (2025) 117487, https://doi.org/10.1016/j.marpolbul.2024.117487.

[8] Y. Liu, B. Han, G. Wang, L. Zheng, Z. Lu, Distribution characteristics, source analysis and ecological risk assessment of polycyclic aromatic hydrocarbons in surface sediments from the western Honghai Bay of China, Mar. Pollut. Bull. 208 (2024) 117001.

[9] Z. Ye, B. Han, G. Wang, X. Lei, S. Qing, Q. Su, X. Chen, Y. Li, Occurrence status, source analysis and risk assessment of polycyclic aromatic hydrocarbons in intertidal surface sediments of typical mangrove wetlands in Guangxi Province, China, Mar. Pollut. Bull. 209 (2024) 117170.

[10] L. Zheng, B. Han, G. Wang, X. Yin, Y. Sun, W. Yan, Polycyclic aromatic hydrocarbons in intertidal surface sediments of mangrove wetlands in Dongzhai Harbor, Haikou, China: understanding sources, distribution, and ecological risk, Mar. Pollut. Bull. 216 (2025) 117950.

[11] Z. Wang, M. Fingas, D.S. Page, Oil spill identification, J. Chromatogr. A 843 (1–2) (1999) 369–411.

[12] K.E. Peters, C.C. Walters, J.M. Moldowan. The Biomarker Guide: Biomarkers and Isotopes in Petroleum Exploration and Earth History, 2nd ed., Cambridge University Press, 2005.

[13] M.B. Yunker, R.W. Macdonald, R. Vingarzan, R.H. Mitchell, D. Goyette, S. Sylvestre, PAHs in the Fraser River Basin: a critical appraisal of PAH ratios as indicators of PAH source and composition, Org. Geochem. 33 (4) (2002) 489–515.

[14] M. Tobiszewski, J. Namieśnik, PAH diagnostic ratios for the identification of pollution emission sources, Environ. Pollut. 162 (2012) 110–119.

[15] E.E. Bray, E.D. Evans, Distribution of n-paraffins as a clue to recognition of source beds, Geochim. Et. Cosmochim. Acta 22 (1) (1961) 2–15.

[16] R.A. Bourbonniere, P.A. Meyers, Sedimentary geolipid records of historical changes in the watersheds and productivities of Lakes Ontario and Erie, Limnol. Oceanogr. 41 (2) (1996) 352–359.

[17] P. Paatero, U. Tapper, Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values, Environmetrics 5 (2) (1994) 111–126.

[18] D.D. Lee, H.S. Seung, Learning the parts of objects by non-negative matrix factorization, Nature 401 (6755) (1999) 788–791.

[19] I.T. Jolliffe, J. Cadima, Principal component analysis: a review and recent developments, Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 374 (2065) (2016), 20150202.

[20] I.C.T. Nisbet, P.K. LaGoy, Toxic equivalency factors for polycyclic aromatic hydrocarbons, Regul. Toxicol. Pharmacol. 16 (3) (1992) 290–300.

[21] E.R. Long, D.D. MacDonald, S.L. Smith, F.D. Calder, Incidence of adverse biological effects within ranges of chemical concentrations in marine and estuarine sediments, Environ. Manag. 19 (1995) 81–97.

[22] D.D. MacDonald, C.G. Ingersoll, T.A. Berger, Development and evaluation of consensus-based sediment quality guidelines for freshwater ecosystems, Arch. Environ. Contam. Toxicol. 39 (2000) 20–31.

[23] Agency for Toxic Substances and Disease Registry, Toxicological Profile for Polycyclic Aromatic Hydrocarbons, U.S. Department of Health and Human Services, 1995.

[24] International Agency for Research on Cancer, Benzo[a]pyrene. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Volume 100F. Chemical Agents and Related Occupations, International Agency for Research on Cancer, 2012.

[25] U.S. Environmental Protection Agency, Method 3545A: Pressurized Fluid Extraction, U.S. Environmental Protection Agency, 1996.

[26] U.S. Environmental Protection Agency, Method 8270D: Semivolatile Organic Compounds by Gas Chromatography/mass Spectrometry, U.S. Environmental Protection Agency, 2007.

[27] K. Zhang, S. Chang, Q. Zhang, Y. Bai, E. Wang, Y. Fan, X. Tu, Q. Fu, L. Wei, et al., Occurrence, source modeling, influencing factors and exposure assessment of polycyclic aromatic hydrocarbons in water sources: a mega-study from mainland China, Environ. Technol. & Innov. 35 (2024) 103634, https://doi.org/10.1016/j.eti.2024.103634.

[28] K. Zhang, A. Qadeer, S. Chang, X. Tu, H. Shang, M.A. Khan, Y. Zhu, Q. Fu, Y. Yu, Y. Feng, Short-chain PFASs dominance and their environmental transport dynamics in urban water systems: Insights from multimedia transport analysis and human exposure risk, Environ. Int. 202 (2025) 109602, https://doi.org/10.1016/j.envint.2025.109602.

[29] B. Barhoumi, K. LeMenach, M.H. Devier, W. Ben Ameur, H. Etcheber, H. Budzinski, J. Cachot, M.R. Driss, Polycyclic aromatic hydrocarbons (PAHs) in surface sediments from the Bizerte Lagoon, Tunisia: levels, sources, and toxicological significance, Environ. Monit. Assess. 186 (2014) 2653–2669, https://doi.org/10.1007/s10661-013-3569-5.