Contents lists available at ScienceDirect

Social Sciences & Humanities Open

journal homepage: www.sciencedirect.com/journal/social-sciences-and-humanities-open

Regular Article

Towards a working definition of coastal community

Ashley Mills *, Salmaan Ansari, Maja Nikšić, Lindsay Forbes

Centre for Health Services Studies (CHSS), University of Kent, Canterbury, UK

A B S T R A C T

The term "coastal community" is widely used in research, policy, and media discourse, yet remains conceptually ill-defined, inconsistently applied, and multivalent. The notion of coastalness often supports definitive - and sometimes contradictory - claims about community health, and socio-economic conditions, particularly when contrasted with "inland" areas. We argue that establishing a clear, consistent, and widely understood definition of coastal and inland communities is a critical precursor to robust, comparable research. Such clarity ensures that comparisons are methodologically sound and that resulting insights can meaningfully inform policy and practice. Drawing on a conceptual analysis of existing definitions alongside a structured assessment of available geospatial boundary data, we discuss this issue in the context of the existing body of research and the problems involved in precisely defining "coastal community". We then propose a working definition of coastalness with respect to towns that lends itself toward meaningful comparative analysis. Our approach defines community as a bounded, named settlement and introduces a continuous, population-weighted measure of coastalness based on distance decay from the coastline. Finally, we discuss the limitations on generalisation of this approach to other settlements and outline future directions for research and policy engagement.

1. Introduction

Physicians in the early nineteenth century promoted hydrotherapy as medicine. The concept of open-air bathing followed, and as British bathing culture developed, patients were encouraged to convalesce or "take airs" at locations around the British coast (Valen, 2022).

We no longer treat Tuberculosis with a trip to the Royal Sea Bathing Infirmary at Margate (Gauvain, 1935), but a trip to the Tate Modern on the Margate seafront is still a popular destination for those looking to escape the humors of London.

The perception that coastal living is "good for you" has permeated modern culture: a Google search for "coastal benefits of living by the sea" reveals a glut of grey literature from wellness blogs, pop science websites, and a plethora of real estate promotional brochures. The perception is amplified in the popular press with articles extolling the virtues of coastal living and some of these claiming to be scientifically informed (Hunt, 2019; O'Grady, 2016).

In 2021, the UK Chief Medical Officer's report into the health of coastal communities (DHSC, 2021) claimed that coastal communities *"have some of the worst health outcomes in England, with low life expectancy and high rates of many major diseases"* (p2) and warns that inaction will lead to *"a long tail of preventable ill health which will get worse"* (p2).

If coastal living is widely portrayed as beneficial to health and wellbeing, as often suggested by the popular press/media, why has the Chief Medical Officer highlighted these areas as facing some of the poorest health outcomes? This prompts a critical examination of what the research reveals.

Unfortunately, as that report points out *"The paucity of granular data and actionable research into the health needs of coastal communities is striking."* (p22, ibid). Specifically, there are a paucity of data indicating level of health in such communities that are granular enough to allow epidemiological analysis of the residents' health.

The aim of this paper is to address this definitional and methodological gap. Specifically, we seek to establish a clear, operational definition of "coastal community" that is suitable for robust comparative analysis in public health research. We do this by critically examining how coastal communities have been defined in existing work, identifying appropriate geospatial units for representing settlements, and proposing a transparent, reproducible measure of "coastalness" at the level of named places. Importantly, we do not begin from the assumption that systematic health disparities between coastal and inland communities necessarily exist; rather, our objective is to define these concepts rigorously so that such questions can be examined empirically.

2. Research context

This section outlines how coastal exposure has been conceptualised and operationalised in influential public health research. We draw on a small number of well-cited studies to characterise the dominant ways in which "coastal" has been defined and measured in this literature.

Our focus is on two methodological dimensions that shape what kinds of spatial and social processes can be examined: the spatial unit at

* Corresponding author. Centre for Health Services Studies (CHSS), University of Kent Canterbury, CT2 7NZ, UK.
E-mail address: ajsm@kent.ac.uk (A. Mills).

https://doi.org/10.1016/j.ssaho.2026.102649
Received 17 September 2025; Received in revised form 26 February 2026; Accepted 5 March 2026

which coastal exposure is defined, and the type of health outcome considered. The examples discussed are used to illustrate recurring operational patterns in the field.

As shown in Table 1, the majority of studies operationalise coastal exposure as distance to the coast, most commonly measured at the Lower Layer Super Output Area (LSOA) level using population-weighted centroids and related to self-reported health outcomes. Coastal proximity is typically discretised into distance bands, although the choice of thresholds varies considerably across studies.

(Asthana & Gibson, 2022) adopt a distinct operational approach, applying a categorical definition of coastal exposure rather than modelling distance gradients. Their analysis draws primarily on objective, GP-recorded morbidity data from the Quality and Outcomes Framework (QOF), supplemented by survey and administrative sources, rather than relying solely on self-reported health outcomes. This classification is applied at the LSOA level, identifying only those LSOAs that overlap built-up areas within 500 m of the mean high water mark; inland LSOAs within the same settlements are not included.

A further variation is provided by (Geiger et al., 2023), who shift the unit of analysis from areas to individual respondents, using self-reported residential distance to the coast (and visit frequency) rather than area-level proximity measures. Despite this shift in scale, distance to the coast remains the underlying exposure construct.

Across these studies, coastal exposure is therefore most commonly conceptualised in terms of distance to the coastline, with the LSOA as the dominant spatial unit and self-reported health as the predominant outcome. Notably, there is little work that treats the town or community as the primary unit of analysis, despite the prominence of these units in policy and public discourse on coastal disadvantage.

A related issue concerns how the coastline itself is defined in distance-based measures. Several studies follow (Wheeler et al., 2012) in calculating distance to the coast using GIS software and a mapped "coastline", without fully specifying how the coastline is defined in complex settings (Wheeler et al., 2012). note that estuaries present practical difficulties for distance calculations and address this by applying an explicit cut-off, defining the end of the coastline at the point where an estuary narrows to approximately 1 km in width. Subsequent studies frequently reference this approach but do not always restate or operationalise this definition in detail, introducing some ambiguity into how distance to coast is measured across the literature.

Table 1
Representative examples of how coastal proximity has been operationalised in public health research. Acronyms: BHPS (British Household Panel Survey), MENE (Monitor of Engagement with the Natural Environment), HSE (Health Survey for England), MHW (Mean High Water Mark), PWC (Population Weighted Centroid).

Study	Outcome measure	Spatial unit	Distance to coast defined as	Coastal proximity categories (km)
Wheeler et al. (2012)	Self-reported health (2001 Census)	LSOA	PWC → MHW	0–1, 1–5, 5–20, 20–50, >50
(White et al., 2013)	Self-reported health (BHPS survey, 1991 - 2008)	LSOA	PWC → MHW	0–5, 5–50, >50
(White et al., 2014)	Self-reported physical activity (MENE survey, 2011)	LSOA	PWC → MHW	0–1, 1–5, 5–20, >20
(Garrett et al., 2019)	Self-reported mental health (HSE, 2011)	LSOA	PWC → MHW	0–1, 1–5, 5–20, 20–50, >50
Asthana and Gibson (2022)	GP-recorded disease prevalence (QOF, 2014/15–2018/19); GP Patient Survey; hospital admissions	LSOA	PWC → MHW	Binary (500m threshold)
Geiger et al. (2023)	Self-reported health and coastal visits	Individual	Residential address → coastline	0–1, 1–2, 2–5, 5–10, 10–20, 20–50, 50–100, >100

2.1. Limits of existing operationalisations

It is difficult to draw meaningful conclusions from these papers because LSOAs themselves are not based on naturally occurring or socially significant divisions. LSOAs are statistical output areas for census taking (Office for National Statistics, 2023). Each contains approximately 1500 people, and they vary in size and shape: populated urban areas are characterised by a high density of small area LSOAs and less populated rural areas have fewer larger-area LSOAs. Therefore, the concept of LSOA is only internally consistent with respect to an abstract space and the decisions that defined the abstraction. This may be in discord with what we want to examine.

More importantly, aggregating LSOAs into arbitrarily wide strips as a function of coastal proximity, across an entire coastline, does not create groupings that map onto meaningful conceptualisations of community. Indeed, the opposite could be argued: that any such strip of land will arbitrarily split coherent places such as cities, towns, and villages into different groups of LSOAs destroying these more meaningful concepts of spatial community whilst simultaneously conflating different types of environments. An example of this is shown in Fig. 1 for Ramsgate, where a 1 km inland boundary is shown to bisect the built up area (BUA) for Ramsgate and its constituent LSOAs.

The Chief Medical Officer's report (DHSC, 2021) emphasises the lack of both *"a nationally agreed definition or consensus on what constitutes a 'coastal community'"* (p20), and explicitly calls to *"Improve data and research into coastal communities"* (p24).

This speaks to a kind of circularity of reasoning. The title of that report *"Chief Medical Officer's annual report 2021: health in coastal communities"*, refers to *"coastal communities"* as if this means something tangible, whilst admitting that there is no consensus definition. Furthermore, the report states that coastal communities *"have some of the worst health outcomes in England, with low life expectancy and high rates of many major diseases"* (p2), which as a statement, if we accept the conclusions of the same report, is both ill-defined and lacking an evidence base to inform it.

Taking this argument to its logical conclusion: is it not presumptuous to assume there is a health disparity at all given the paucity of existing research and nebulous conceptualisation of health and place? On what basis and between what is such a disparity presumed to exist?

If we don't clearly define what we mean by "coastal community", or measure it in a way that reflects lived realities, we risk misreading the true nature of health inequalities across geographical space.

Therefore, we should not begin with the assumption that disparities between coastal and inland communities exist and then define these terms to fit the assumption, on the contrary, we should define these terms rigorously and then examine the data to see if disparities exist.

The aim of this paper is therefore not to assume the existence of coastal health inequalities, but to lay the groundwork for a practical, working definition of geospatial community and its coastalness, capable of supporting robust comparative analysis in future scientific and policy-oriented research.

3. Methods

Building on the definitions set out above, this section describes how we operationalise community and coastalness using existing geospatial data. Our aim is not to discover a single "true" definition of coastal community, but to construct a transparent, reproducible measurement framework that aligns with the conceptual requirements identified

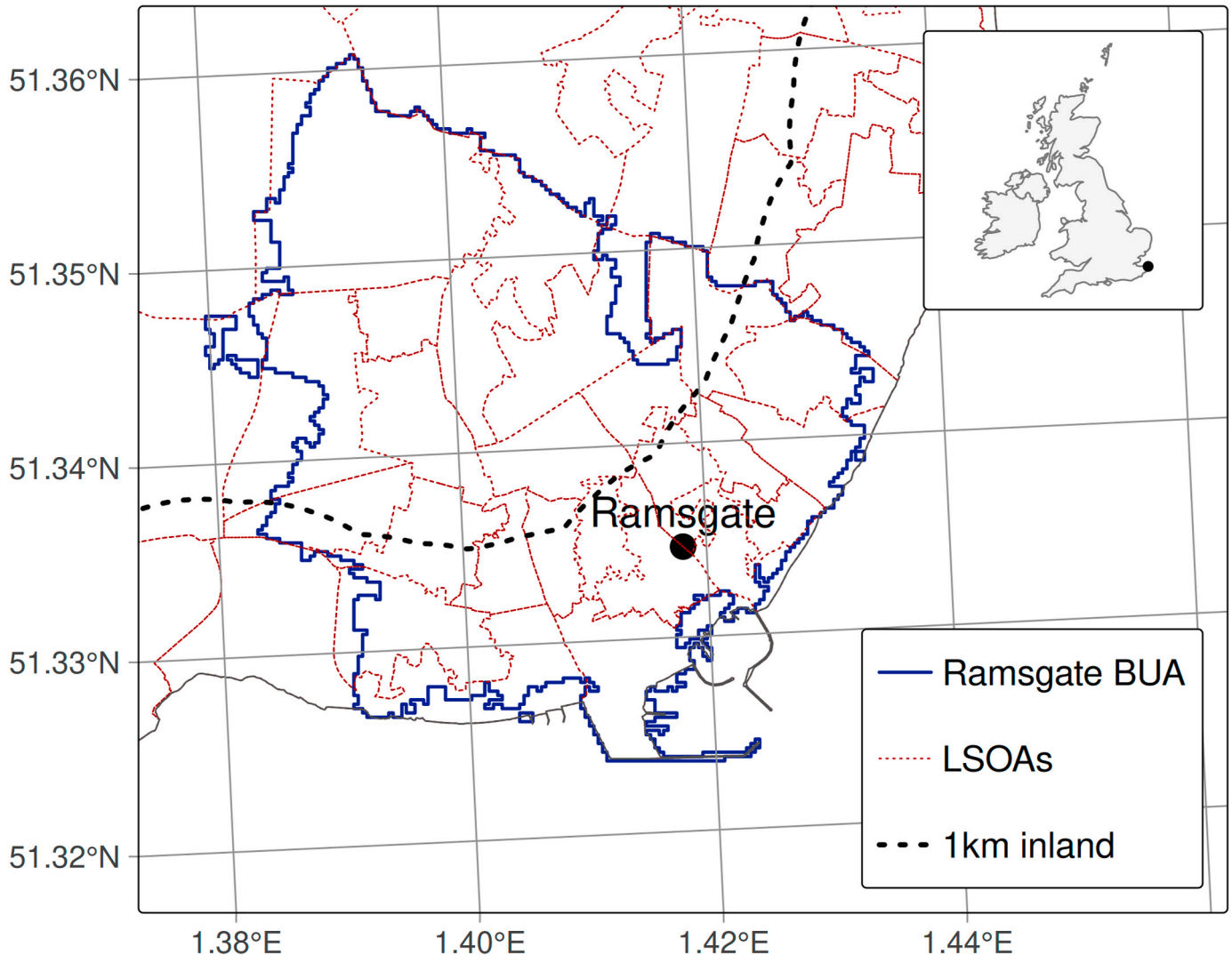

Fig. 1. The effect of splitting Ramsgate BUA by 1 km inland boundary. Plot uses 2024 BUAs and 2021 LSOA boundaries (latest release at time of analysis).

earlier: spatial coherence, population relevance, and graded proximity to the coast. We therefore proceed in four steps: (i) identifying candidate boundary datasets capable of representing settlements, (ii) selecting appropriate boundary proxies for different settlement types, (iii) defining a practical and reproducible representation of the coastline, and (iv) computing a continuous, population-weighted measure of coastalness for bounded communities.

3.1. Defining community and coastalness

3.1.1. Community as a geospatial unit

An Office for National Statistics terminology guide states that "'Community' is a very general term referring to the people living in a locality or to the locality itself." ((Office for National Statistics, 2023) p75). More broadly, community is also used to refer to demographic subsets of a given locality, ethnic diaspora, or even arbitrary associations and groups with a shared sense of belonging (Jewkes & Murcott, 1996; MacQueen et al., 2001). To measure it effectively, the term "community" must be narrowed in scope and clearly defined.

In the introduction above we used the Chief Medical Officer's report into the health in coastal communities to show that using ill-defined terms to make definite statements leads to circular reasoning. But the bulk of that report stands in contrast to this tautological stance since most of the material is neither abstract nor nebulous: rather the report documents a series of case studies of real towns and cities situated on the coast. It looks at real places and makes measurements about real places, and then arrives at conclusions about their differences.

What we can deduce from looking at this aspect of the report is that the authors have already accreted a concept of "community" that implies geospatial coherence and association with a name. Geospatial coherence further implies the existence of boundaries to separate entities from each other.

We argue that this is the natural unit of study for this domain since "coastal" and "inland" are geospatial terms and named place is the smallest unit of geospatial coherence which is meaningfully integrated into local administrative structures (councils, healthcare provision, etc). Thus we define community in this context as:

Community: *a geospatial bounded area encompassing the residents of a single named place.*

Reducing community to an exclusively geospatial definition shatters the multivalent complexity of the original overloaded term. This allows us to separate out factors that are adjunct to the broader meaning of community, such as sense of belonging, but which are better measured at the level of the individual. Community is a word that has lots of different meanings layered on top of each other; here we isolate the spatial components and relocate the others to their more natural measurement abstractions and frames of interpretation.

3.1.2. Coastalness as a property of communities

It is implicitly assumed in coastal health research that health inequalities are caused by measurable differences between communities and that the "coastalness" of a place is both a measurable and meaningful abstraction with respect to health outcomes.

It would however be a mistake to directly treat "coastalness" as an actual determinant of health. If there are any health disparities at the level of coastal community, then it is likely to be because geospatial "coastalness" is a superposition of latent variables which do have bearing on health, perhaps for example: disparities in access to opportunities, transport, ozone levels, sun exposure, cultural expectation, etc. Note that some of these are still secondary or tertiary and require further decomposition to reach their causal forms.

This might seem like a glib point to make but it is important: meaningful geospatial differences exist such as "living above granite" which demonstrate a material impact on health when measured, but "measured exposure to Radon" is closer to causality. Similarly: if coastalness implies disparities, there will be mechanisms underneath it that are closer to causality.

It is also possible that coastalness at the level of geospatial community is not meaningful, or hides too much latent variance to be meaningful and shows no disparity. Therefore we will not make an assumption that coastalness is a meaningful abstraction at the level of geospatial community, rather our approach here is to discover whether it is.

In summary, in order to analyse disparities at the level of community: which in practice means city, town, village, hamlet, or other settlement we need to precisely define these concepts so we can aggregate and analyse statistics at these abstractions in order to understand and treat the responsible determinants. To do this we need boundaries for communities attached to each of the named places.

3.1.3. Implications for measurement

Defining community as a bounded, named settlement and coastalness as a property of that settlement has immediate implications for how coastal exposure should be measured.

First, it implies that the unit of analysis must be a spatially coherent area rather than an abstract statistical zone or an arbitrary distance band. If the object of interest is a community, then exposure cannot be meaningfully defined solely at the level of points (e.g. centroids) without reference to the geometry and population distribution of the area being studied. Measures that assign coastal status to communities based on a single representative point risk mischaracterising places whose populations are spatially heterogeneous or whose built form interacts unevenly with the coast.

Second, treating coastalness as a continuous property rather than a binary label implies that measurement strategies should preserve gradation rather than impose hard thresholds *a priori*. Binary classifications collapse potentially informative variation and embed implicit assumptions about where the coast "ends." If coastalness operates as a proxy for multiple underlying mechanisms, then the strength of that proxy is likely to vary smoothly with distance and spatial configuration rather than discontinuously at an arbitrary boundary.

Third, this framing requires that population distribution be explicitly incorporated into measurement. Communities are not uniformly populated across their spatial extent, and coastal influence on residents is mediated by where people actually live within a settlement. Measures that ignore population weighting implicitly assume uniform exposure, an assumption that is rarely justified in practice.

Finally, these definitions imply that measurement must be transparent about its assumptions and tunable parameters. Choices about distance decay, coastline definition, and boundary inclusion are not neutral; they encode judgements about how space, proximity, and influence operate. Rather than concealing these choices within opaque heuristics or fixed thresholds, they should be made explicit so that their consequences can be examined, challenged, and adjusted.

Taken together, these considerations motivate a measurement approach that (i) operates over bounded settlement areas, (ii) incorporates spatial geometry and population structure, and (iii) represents coastalness as a continuous, interpretable quantity. In the sections that follow, we describe how these principles are implemented using existing public geospatial data.

3.2. Identifying candidate boundary data

We would like to be able to measure relevant properties of every geospatial community (as defined in the last section) in the UK, including coastalness, and relate these to health outcomes. However, to aggregate statistics for a community, it needs to be bounded.

Ordnance Survey (OS) is the UK's national mapping service, it is both a government agency and public corporation but serves as the authoritative source for geographical entities. According to OS nomenclature, there are four broadly recognised geospatial communities in the UK: city, town, village, and hamlet. These are referred to collectively as different types of settlement: a place where people settle.

UK settlement categories (city, town, village, hamlet) are not defined using a single, consistent set of quantitative criteria. City status is conferred by Royal Charter and is therefore historically designated rather than determined by size, population, or function. Other settlement types are labelled using a combination of cartographic convention, historical usage, and relational comparison, rather than explicit thresholds. While some planning and descriptive sources refer to indicative characteristics such as the presence of a commercial centre or the extent of built development, these features are not codified as formal decision rules within Ordnance Survey settlement datasets. Villages and hamlets, in particular, are defined primarily in relation to larger settlement types, without fixed population or area thresholds. Despite this conceptual imprecision, these categories constitute the only nationally standardised settlement labels available and therefore serve as a pragmatic reference framework for settlement-level analysis.

To illustrate the boundary problem let us use "Town" as an example of a community for which we want to aggregate statistics. OS maintains a list of settlement names, their type, and their centroids in a product called OS Open Names that is free to use (Ordnance Survey GB, 2023b).

These settlement centroids are not population derived but for cities and towns are determined by a type of common knowledge heuristic: the technical specification states that town and city centroids represent *"the position that the majority of informed people would accept as being the 'centre' of the settlement"* (Ordnance Survey GB, 2023a) (p4). And for smaller settlements *"the position has been generated from the major road junction within the settlement"* (ibid p4).

Aside from the observation that these town centroids are heuristically defined and do not necessarily reflect settlement population distributions accurately, the centroids are nevertheless available: we can put names to places. This seemingly obvious statement contrasts starkly with that of settlement extents: **there are no canonical boundary data for UK settlements.**

Without boundaries, we cannot reliably compare health outcomes between settlements, estimate exposure to environmental risks, or allocate resources based on spatial need. We are left measuring the health of populations without knowing their shape or scale of measurement.

While there are no canonical boundaries for UK settlements, there are other defined boundaries, with published extents that might serve as proxies.

The Office for National Statistics (ONS) maintains the Open Geography Portal (ONS, 2025c); a tool which can be used to explore all of the major published boundaries used for UK government administration and statistics, plot different boundaries on a map of the UK, and download these boundaries in common geospatial encodings to enable local analysis. As of July 2025, the portal contains 59 boundary datasets in 7 categories and 12 centroid datasets.

To aggregate statistics by settlement, we need boundaries which meaningfully capture population structure, are aligned with named places, and remain stable enough for longitudinal analysis.

We have qualitatively assessed the suitability of these 59 boundaries for representing geospatial communities using the following criteria.

1. Do the boundaries map cleanly onto named settlements?
2. Are they spatially coherent and population-representative?
3. Are they consistent enough for comparison and reproducibility?

The results are summarised in Table 2.

From this table we have concluded that Parish and BUA are the most viable candidates to serve as proxies for settlement extents. This does not mean that the other boundaries are not useful, for example: it may be useful to include the enclosing healthcare trust as a covariate in analysis, but we do not consider the excluded boundaries useful for defining settlement extents.

In the following we have restricted the scope of our analysis to England and Wales. This is due to the fact that Scotland's geospatial and settlement data are managed independently by the National Records of Scotland, with different formats and definitions that would require a separate harmonisation process.

3.3. Selecting boundary proxies for settlements

Because no canonical boundaries exist for UK settlements, any settlement-level analysis requires the use of proxy boundaries. From the candidate datasets identified above, we therefore sought to determine which boundary types most plausibly represent the spatial extent of settlements for analytical purposes. Our objective at this stage is not to assert that any proxy is a perfect representation of a community, but to evaluate available options and select those that provide the most consistent and interpretable mapping between named settlements and bounded populations.

Table 2
Suitability of ONS boundaries for representing the populations of settlements. Acronyms: Lower Layer Super Output Area (LSOA), Integrated Care System (ICS).

Category	N	Example Boundaries	Evaluation	Candidate
Administrative	13	Country, Local Authority, County, Parish	Most are too large and have one-to-many relationships with settlements. Parishes don't fully cover population, but are one-to-one for some settlements and likely to be population coherent.	Parish
Census	9	LSOA, Built Up Area (BUA)	LSOAs and larger abstractions are statistical units, not place-aligned. BUAs attempt to define contiguous built-up land associated with named settlements, are one-to-one for some settlements.	BUA
Electoral	7	Wards, Divisions	Frequently change, designed for equal representation, often split or merge settlements unnaturally.	None
Grid	1	1 km Grid Squares	Arbitrary spatial tiling, no relation to population or place identity.	None
OECD/ Eurostat	8	Territorial Level Regions	Supra-national classification, not relevant to UK settlements.	None
Health	11	ICS, Cancer Alliances, Trust Boundaries	Spatial units too large and not designed to represent discrete settlements.	None
Other	10	National Parks, Police Force Areas	Thematic or environmental, not linked to settlements.	None

3.3.1. Built up areas (BUAs)

Built-Up Areas (BUAs) are spatial extents produced by Ordnance Survey to represent contiguous areas of built development.

According to the Ordnance Survey technical specification for BUAs (Ordnance Survey, 2022) they are defined on top of the National Land Use Database (NLUD) which was/is UK government project to map land use in the UK for the primary purpose of planning. The last release was in 2006 and is known as NLUD 4.4 (Office of the Deputy Prime Minister, 2006). The output was a land use classification map at a resolution of 25m × 25m squares. It isn't clear from the documentation exactly how this output was arrived at, as prior versions of NLUD are rooted in origins going as far back as 1980, but a mixture of fieldwork validation and examination of OS Mastermap are mentioned as primary tools.

The BUA technical specification states that cells with NLUD types "Building, Road, Residential Garden, Made Surface, or built-up type" were considered built up, and then assigned names according to OS open names centroids for settlements. It isn't clear how this process was achieved, whether a nearest settlement approach was used or some other method.

After each cell is given a settlement name, adjacent cells with the same name are merged, iteratively, until no more merging can occur. Then cells with a total area of 20,000 m^2 or more are kept as built up extents. Built up areas are then derived by excluding any water bodies or non "built-up" cells that might exist inside them.

While the documentation does not provide a detailed empirical justification for the specific value of 20,000 m^2, the threshold functions as a pragmatic filter to ensure that BUAs represent settlements of sufficient spatial coherence to be meaningful for statistical reporting. An important consequence of this design choice is that smaller settlements, particularly hamlets, are less likely to be represented as BUAs, a limitation we account for explicitly in our subsequent boundary evaluation.

The BUAs are then named in a process that uses the *"the most significant constituent settlements"* (whatever that means) and arbitrates duplicates using overlying administrative boundaries: local authorities, parishes, etc as a post-fixed name.

OS were contacted for clarification about the ambiguities outlined here but have yet to respond with an answer.

This ambiguity in the definition of BUA is not ideal, but it does give us a reasonable bounded area that is tied to a settlement and its resident population. An example was already shown for Ramsgate in Fig. 1.

3.3.2. Parishes

Parishes began as church catchment areas, but evolved into civil parishes, which today form the minimal administrative unit for a settlement or group of settlements. Their boundaries are not fixed: they can be redrawn, merged, or even abolished, and are sometimes shaped by political or practical considerations rather than community identity.

Parishes are included here as candidate boundary proxies because they are widely used in official statistics and often correspond to named places, making them a plausible unit for aggregating population-level data. However, unlike BUAs, parish boundaries are not derived from the spatial extent of built development and may encompass large areas where people do not live, multiple settlements or subdivide a single settlement. As a result, their suitability as proxies for settlement extents is an empirical question, which we assess explicitly in the boundary coverage analysis below.

3.3.3. Boundary coverage assessment: design

BUA and parish are our candidates for settlement boundary proxies

but their utility will depend on the uniqueness of their relationships to settlement centroids.

3.3.3.1. Aim. To determine whether Built-Up Areas (BUAs) or Parishes provide the more reliable proxy for representing settlements in our analysis.

3.3.3.2. Method. For each settlement centroid in England and Wales, we tested whether it falls within a BUA or a parish boundary. Where it does, we then assessed whether the relationship is one-to-one (the centroid maps uniquely to a single boundary) or shared (multiple settlements fall within the same boundary). We used the latest release of each dataset at the point of analysis: the 2024 release of ONS Open Names, 2024 BUA boundaries, and the May 2023 Parish boundaries coupled with the 2024 boundries for "Non Civil Parished Areas" as the parishes are split across two separate datasets.

When relationships are not one-to-one, we account for settlement hierarchy: if a BUA contains both town and village centroids, it is reasonable to treat the villages as subsumed into the town, and still take the BUA as the valid representation of the town's population. By contrast, a BUA containing both a large town and a smaller village cannot be said to represent the smaller village.

We applied this approach to settlements classified as City, Town, Village, and Hamlet. We excluded ONS's "Other Settlement," as this category mixes administrative units (e.g. London boroughs) with other settlement types and therefore introduces ambiguity. The outcomes of this assessment are reported in the Results section.

Based on the boundary coverage assessment reported in the Results section, we adopt Built-Up Areas (BUAs) as the primary proxy for bounding towns in subsequent analyses. BUAs provide the most consistent and interpretable mapping between named towns and spatially coherent populations, particularly when allowing town boundaries to subsume smaller settlements. Parishes are retained as a secondary proxy for hamlets, where BUA coverage is sparse. Having established an operational representation of settlement extents, we next address how the coastline itself is defined for the purposes of measuring coastal proximity.

3.4. Defining the coastline

To know if something is coastal we need to know what the coast is. If we try and define the coast as the interface between land and sea, we are frustrated by the fact that this interface changes due to tides. It appears that this issue has not been clearly addressed or defined to give a canonical "coastal boundary".

ONS provide two boundaries for country extents: those "clipped to the coastline" and those which are the "extent of the realm" which ONS notes *"(usually this is the Mean Low Water mark but, in some cases, boundaries extend beyond this to include offshore islands)"* (ONS, 2025a). Both boundaries are candidates for the coastline, but both have issues. Fig. 2 provides an illustration for an area of Kent, UK.

The "extent of the realm" is shown as the exterior dotted red line. It cuts off the southern extent of the Isle of Sheppey from being "coastal" along with the northern coast of Kent from Gillingham to Faversham. It also places Canvey Island "inland" due to where the Thames is clipped. The extent of the realm boundary also biases settlements with large tidal ranges towards being classified as non-coastal. For example Cliffsend (not shown in figure) is a coastal settlement where the tide, and hence the "extent of the realm" is ~1 km away from residential areas whereas "clipped to the coast" is only ~200m.

The "clipped to coast" boundary is shown in Fig. 2 as the interior solid blue line and suffers from the issue of ingressing too far inland due to the tidal nature of rivers, for example: The Medway's tidal nature places Alyesford as "inland" despite it being ~20 km to the nearest

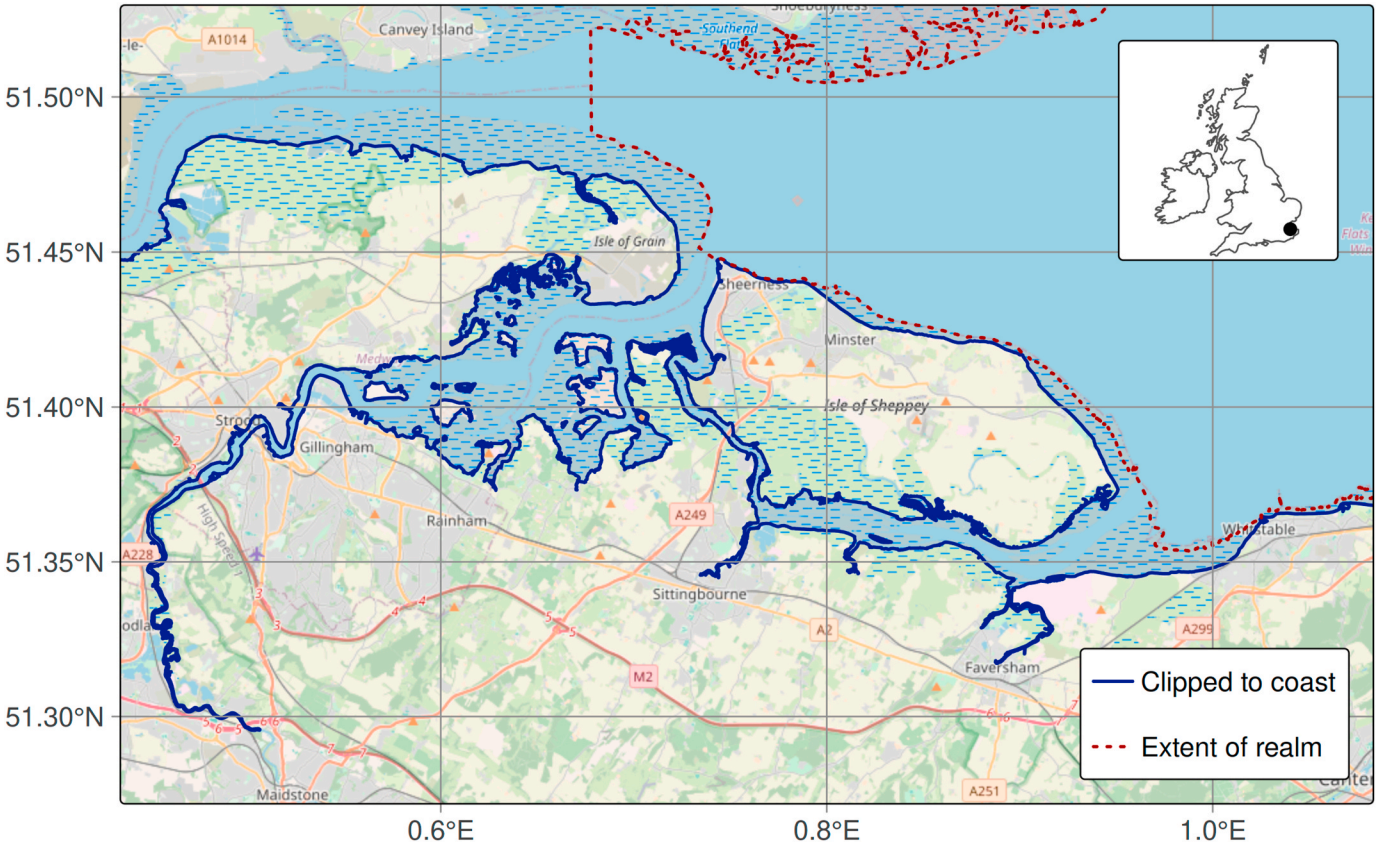

Fig. 2. Contrasting "clipped to the coast" boundary (solid blue line) with the "extent of the realm" boundary (dotted red line).

"coast".

ONS does provide a third definintion of coast: ONS have defined the shoreline for census reporting on Built up Areas (ONS, 2024) and provide a GitHub repository (ONS, 2025b) where they describe how the shoreline is constructed: *"The ONS GB Shoreline is created by taking full resolution extent of the realm boundaries (also known as low water mark) and 'cookie-cutting' out selective areas of sea, tidal water and tidal land taken from Ordnance Survey data."*.

Unfortunately the process depends on commercial data and thus is not easy to replicate, neither is it obvious why or how "selective areas" have been removed. We do nethertheless encourage ONS to release the pre-computed "GB shoreline" boundary for researchers to use.

In its absence, we have noted that the ordinary "clipped to the coast" boundary, which includes tidal rivers, would be ideal if the tidal rivers were sensibly clipped since it would: (i) Not suffer from the "extent of the realm" ambiguities described above (ii) be derivable from simple open boundary datasets.

It is of course a subjective question as to where to clip tidal rivers, and how far up an estuary is "coastal", but these questions must be answered if we are to arrive at a sensible definition of the "shoreline" or "coast". The only clear precedent in the coastal health literature appears to be (Wheeler et al., 2012), which after noting the tidal river problem adopts an arbitrary cut-off of 1 km which subsequent research adopts by citation.

We leave the general problem of defining the estuarine coastline to others. For our purposes in Kent, only six tidal rivers or channels appear in the ONS vector data intersecting the ONS coastal boundary: we examined them individually. The Stour and the Swale are each ~100m wide at their seaward limits, and the channels at Conyer and Sittingbourne are ~150m wide; in all four cases we cut them at their mouths, as no meaningful ambiguity exists. The two remaining cases required local judgement. We clipped the Medway at Chatham docks, reflecting its established role as a functional dockyard, and the Thames at Gravesend, which we treated as the last settlement plausibly coastal in character, given its strong maritime history and the point at which the estuary narrows sharply into river.

It seems likely that across the rest of the UK most tidal rivers could be clipped at the seaward limit of their mouths, while the larger estuaries would require local knowledge and therefore some degree of subjectivity, to decide where the "coast" should end. This judgement need only be made once, with subsequent updates required only if new settlements are constructed or significant sea-level changes occur.

3.5. Deciding which bounded areas are "coastal"

Now that we have (i) discovered that there exist boundaries that are reasonable proxies for settlements and (ii) decided what the coast is, we can ask which of these boundaries are coastal and which are not?

Logically it follows from the fact that the coast is a geographic boundary, that spatial proximity will be involved in our definition. Indeed, we saw earlier that many prior investigations into "coastal health" used aggregations of LSOAs within specific distances of the mean high water mark as a dividing boundary. Most of these use LSOA population weighted centroid as the measuring point, but a boundary is not a point: it encloses a space of points.

If we refocus the concept of distance through the lens of our more meaningful boundaries: BUAs, we can see that as boundaries grow in size, they become less point-like and their geometry becomes more and more important. It is not meaningful to speak of the distance between a shape and the coast without being clear what we mean by this. How exactly should we measure distance to coast over a shape and arrive at a clear categorisation of coastal/inland? Do we treat an elongated boundary where only a small part of it is "near" to the coast the same as a boundary that is more aligned with the coast?

We saw earlier that much of the existing research uses LSOA as the unit of analysis and assigns coastal status based on a binary or categorical distance threshold to the coast. This approach embeds no concept of geometry, and uses heuristically obtained distance thresholds.

One of the more useful strides towards classifying coastal communities comes from a recent ONS publication where BUAs were classified as coastal or not, and comparative statistics were reported for basic demographics and self-reported health (ONS, 2024).

The criteria adopted by ONS was to classify a BUA as coastal if at least one of the following conditions were met.

1. BUAs which have a boundary within 1 km (km) of the statistical coastline and with a surface area of 50% or more within 3 km of the statistical coastline
2. BUAs which have a perimeter of more than 2.5 km within 25 m of the statistical coastline

Note that "statistical coastline" here means the ONS heuristic defined for the purpose of the report (ONS, 2025b) which we discussed earlier as the GB shoreline.

ONS do not explain how the two criteria above were arrived at and whether or how the methodology has been validated for consistency with human understanding. The methodology documentation does state that they had to manually re-label Middlesborough as non-coastal. This implies that local knowledge was used to bootstrap and define the methodology but that the final classification system failed in at least one case.

A general issue with any binary classification system based on geographical constraints is that the constraints end up being adjusted to fit the *a priori* distribution, which can result in overfitting and loss of generality, for example: will the ONS system correctly classify all villages on the coast that become towns in the future? Will it re-classify some towns as non-coastal if they extend too far inland?

In the ONS case for BUAs discussed above it is reasonable to assume that the definition was heuristically derived based on *a priori* labelling of towns.

Others have attempted to empirically derive a definition of coastalness based on model parsimony with respect to a given health outcome. For example in (Welsh et al., 2024) the authors derive a coastal proximity indicator (CPI) for cancer research by constructing a pool of candidate variables that encode various coastal parameters. They then rank these candidates according to model Bayesian information criterion, after first filtering out any CPIs that don't produce significant associations with cancer prevalence.

Through this process, they conclude that their best model is one they call G_25_5 which translates as "LSOA is coastal if 25% or more of the postcode centroids are within 5 km of the coast". Then, they explain through the lens of this coastalness parameter that *"age-standardised incidence was 9.1 per 100,000 persons higher (95% CIs = 5.6 to 12.6), crude prevalence was 6.4 per 100,000 people higher (95% CIs = 4.4 to 8.4) and age-standardised mortality was 1.6 per 100,000 people higher (95% CIs = −0.8 to 4.0) compared with non-coastal LSOAs."*.

This a clever approach in principle because it attempts to frame coastalness in terms of its utility toward a specific outcome measure, but in practice it has problems.

First, the CPI candidate space is arbitrary, relying on a small set of qualitatively chosen distance thresholds (e.g. 1 km, 3 km, and 5 km) rather than a principled or continuous representation of coastal proximity. Second, a large number of candidate indices are evaluated without correction for multiple testing. Scanning many correlated candidates increases the risk that statistical significance arises by chance rather than reflecting a robust underlying association. Because candidates are then filtered using a $p < 0.05$ criterion, the approach risks double selection, increasing the likelihood that the retained CPI reflects overfitting or residual confounding - such as alignment with age structure or other spatial characteristics - rather than a substantively meaningful measure of coastal exposure.

Finally however, the biggest weakness of this approach is that it uses LSOAs as the unit of analysis: and as we have seen LSOAs are not settlements. And furthermore, the semantic relevance of G_25_5 is unclear, it is not explanatory; it is artificial in the technical sense.

What this really reveals is that we need to think about geography in a more neutral and flexible way.

3.6. Defining coastalness for bounded settlements

We propose a method of classifying bounded areas that captures both geometry and distance, by exponentially weighting LSOAs within any defined boundary by their distance from the coast, and scaling by their population:

Let the LSOAs whose population weighted centroids fall within a given boundary be indexed by i. Then we define the weighted coastal influence *WCI* of a bounded area as:

$$WCI = \sum_i P_i e^{-\lambda d_i}$$

Where P_i is the population of LSOA i, d_i is the crow distance between the population weighted centroid of LSOA i and the coast, and λ is a tunable exponential decay coefficient that determines how rapidly coastal influence wanes.

We can normalise this value using the population sum:

$$Coastalness = \frac{\sum_i P_i e^{-\lambda d_i}}{\sum_i P_i}$$

This gives us a value in $[0, 1]$ that characterises the degree of "coastalness" of a bounded area.

We propose to attune λ using *a priori* local knowledge for a given area, for example Kent, so that it meets intuitive truth at a coastalness threshold of 0.5 such that settlements with a coastalness above 0.5 are considered "coastal" and those below 0.5 are considered "non-coastal".

The purpose of this isn't to assert our understanding onto the system as ground truth, but to establish an informative prior so that the posterior use of the measure is loosely interpretable. This way we avoid binary measures that throw away complexity (we can use the coastalness parameter directly in the model, or as a discretised categorical variable) but we retain interpretability through a binary lens if we want it.

The algorithm works for any bounded area that represents a settlement, but here we focus on BUAs and towns.

We illustrate the proposed definition using towns in Kent. Kent was selected for pragmatic and methodological reasons: the authors are based in the county and therefore have detailed local geographical knowledge, particularly in relation to complex estuarine boundaries and settlement structure. This familiarity allows explicit and transparent decisions to be made where automated or national boundary definitions are ambiguous. Kent also contains a diverse range of coastal and inland towns, making it a suitable test case for demonstrating the behaviour of the coastalness metric. The resulting coastalness values and classifications for towns in Kent are presented in the Results section.

4. Results

4.1. Boundary coverage assessment

4.1.1. BUAs

Table 3 summarises the results of the boundary coverage assessment for Built-Up Areas (BUAs). It reports the extent to which settlements of different types are contained within BUAs and whether those relationships are one-to-one or shared with other settlements.

To explain this table we will describe the numbers in the first row for Type "City". There are 63 city settlement centroids in the ONS data for England and Wales (N). Of these, 63 fall inside a BUA boundary. For 45 cities, that relationship is one-to-one. Two cities: The City of Westminster and the City of London share the same BUA despite being contiguous because historically they were granted city status independently. 16 cities share their BUA with centroids for smaller place types (Towns, Villages, or Hamlets). No cities share their BUA with bigger settlement type since they are the largest type. No cities share their BUAs with both smaller and larger settlement types simultaneously.

From this table we can see that the relationship between Town and BUA is most consistent, with only two towns not being inside a BUA, and 86.44% having a one-to-one relationship with a BUA. There are no towns that share their BUA with another town and town BUAs only subsume smaller settlement types.

For villages, 48.07% of villages are associated with a BUA and 34.25% have a one-to-one relationship with a BUA, 11.54% of villages share their BUA with another village. There are villages that share their BUA with both smaller and larger place types, including 26 simultaneously.

Hamlets are under-respresented in terms of BUAs, since the likelihood of being built-up tends to reduce with settlement size, and the BUA methodology requires a minimum area of 20,000 m^2. Only 4.16% have a BUA, and only 0.83% have a one-to-one relationship with a BUA, the majority sharing a BUA with other settlements of the same or larger type.

Note that the columns to the right of "Has BUA" do not necessarily sum up to "Has BUA" since a settlement can share its BUA with places of the same geospatial type and settlements of different geospatial types simultaneously.

For example, the Westminster-City of London BUA contains two settlements classified as cities, while also subsuming multiple smaller settlements classified as towns or villages. In this case, each city shares its BUA with a settlement of the same geospatial type and with settlements of different geospatial types simultaneously.

One issue encountered when mapping centroids to BUAs is that major roads are often excluded from BUAs due to having a non "built-up" NLUD classification. Settlement centroids however, are sometimes placed on crossroads. This leads to the (fortunately rare) scenario where a settlement centroid is not in the BUA of the same name. An example of this is shown in Fig. 3.

The Figure shows Basildon settlement centroid as a dot, in an area excluded from the Basildon BUA because it includes major roads and a park.

We accounted for this in our membership calculations by manually checking the 13 (~1% of) towns without a BUA assignment and

Table 3
Relationship between ONS settlement centroids and BUA boundaries, by settlement type. It shows how ONS-defined settlement centroids of different settlement types (City, Town, Village, Hamlet) relate to Built-Up Area (BUA) boundaries in England and Wales. For each type, we report the number of centroids (N), how many fall within a BUA, how many have a one-to-one mapping with a BUA, and how many share a BUA with either the same type, smaller types, or both.

Type	N	Has BUA	Unique BUA	Shared Same Type	Shared Smaller Type	Shared Bigger Type	Shared Both Type
City	63	63 (100.00%)	45 (71.43%)	2 (3.17%)	16 (25.40%)	0 (0.00%)	0 (0.00%)
Town	1143	1141 (99.83%)	988 (86.44%)	0 (0.00%)	153 (13.39%)	0 (0.00%)	0 (0.00%)
Village	13,662	6567 (48.07%)	4679 (34.25%)	1576 (11.54%)	350 (2.56%)	130 (0.95%)	26 (0.19%)
Hamlet	11,083	461 (4.16%)	92 (0.83%)	64 (0.58%)	0 (0.00%)	348 (3.14%)	0 (0.00%)

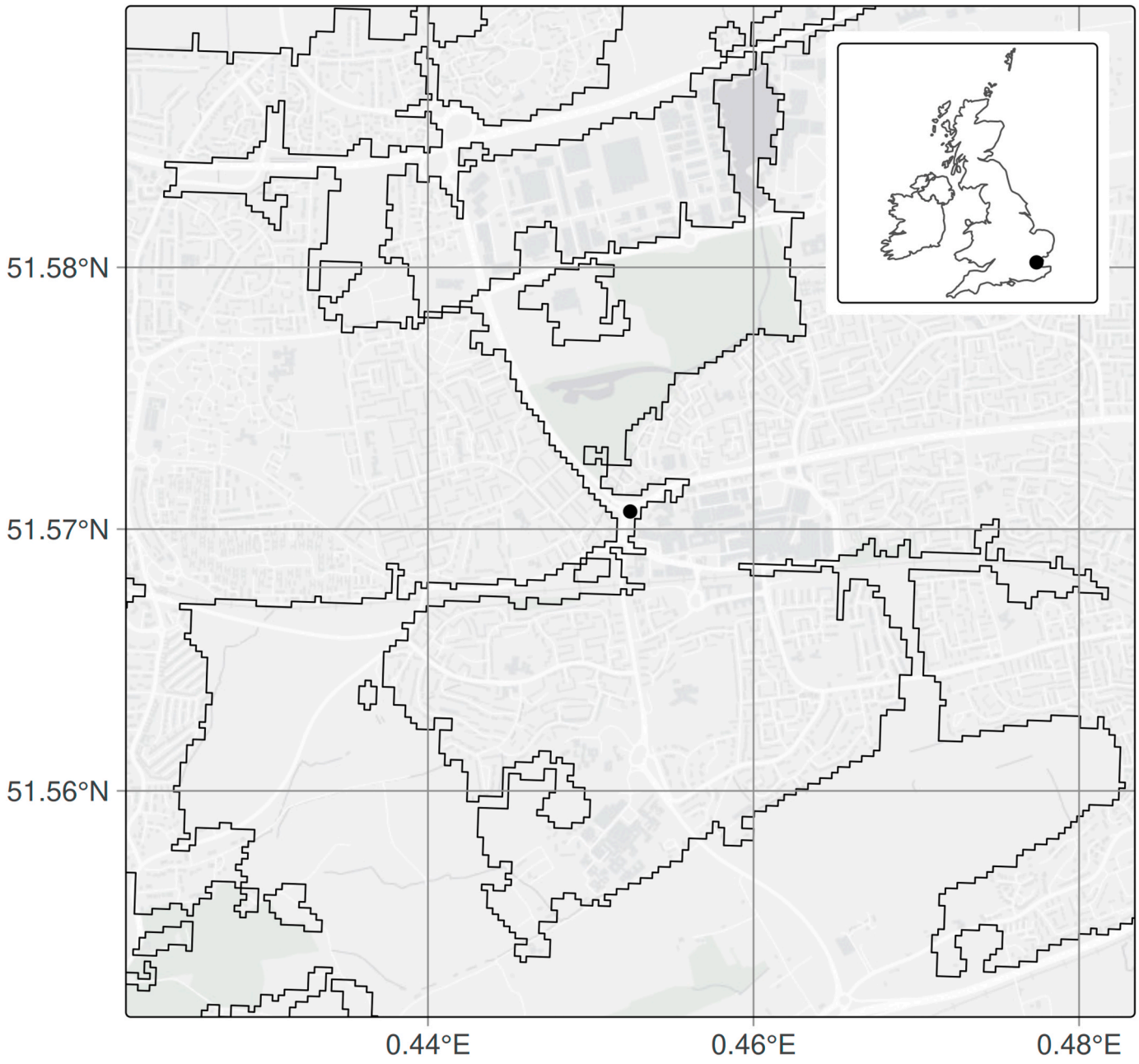

Fig. 3. Basildon settlement centroid (black dot) in relation to Basildon BUA (drawn boundary lines). The settlement centroid falls in an area excluded by the BUA.

assigning them correctly where required. A limitation is that we did not extend this manual check to villages or hamlets, due to the large numbers of these settlements without BUA membership, but we expect this to affect a small proportion of these settlements (<1%) due to the distribution of major roads.

4.1.2. *Parishes*

Table 4 summarises the results of the boundary coverage assessment for parish boundaries. As with the BUA analysis, it reports coverage and uniqueness of centroid–boundary mappings across settlement types, providing a basis for comparing administrative and built-form boundary proxies.

Table 4
Relationship between ONS settlement centroids and Parish boundaries, by settlement type. This shows how ONS-defined settlement centroids of different settlement types (City, Town, Village, Hamlet) relate to Parish boundaries in England and Wales. For each type, we report the number of centroids (N), how many fall within a Parish, how many have a one-to-one mapping with a Parish, and how many share a Parish with either the same type, smaller types, or both.

Type	N	Has Parish	Unique Parish	Shared Same Type	Shared Smaller Type	Shared Bigger Type	Shared Both Type
City	63	17 (26.98%)	11 (17.46%)	0 (0.00%)	6 (9.52%)	0 (0.00%)	0 (0.00%)
Town	1143	850 (74.37%)	399 (34.91%)	20 (1.75%)	447 (39.11%)	0 (0.00%)	0 (0.00%)
Village	13,662	13148 (96.24%)	3500 (25.62%)	7219 (52.84%)	7096 (51.94%)	621 (4.55%)	428 (3.13%)
Hamlet	11,083	10902 (98.37%)	908 (8.19%)	7868 (70.99%)	0 (0.00%)	8767 (79.10%)	0 (0.00%)

The results show that BUA is a much better proxy for Town than Parish is, with only 34.91% towns having a one-to-one relationship with parish. In comparing the suitability of Parish as a proxy for Village against BUA as a proxy we can see that 96.24% of Villages have a Parish in comparison to 48.07% having a BUA. But in practice more Villages have a one-to-one mapping with BUA than parish.

For Hamlets we can see that 8.19% map one-to-one to a Parish in comparison for 0.83% mapping one-to-one with a BUA, and thus Parish may be the most complete proxy for Hamlet despite having low coverage. Whether or not that coverage is useful will depend on the research question.

4.1.3. Summary and boundary selection

In conclusion, the most appropriate candidate for bounding geo-spatial communities is BUA for Towns and Villages, and Parish for Hamlets. Towns have the best BUA coverage at 99.83% of all towns, when we allow town boundaries to subsume villages. For this reason we proceed toward a definition of coastalness for towns, since we can most reasonably bound them using BUAs.

4.2. Coastalness metric

We applied the coastalness algorithm described in the Methods section to towns in Kent. This produced a continuous distribution of coastalness scores across the interval [0,1], with towns conventionally regarded as coastal generally exhibiting higher values and inland towns clustering near zero.

Using $\lambda = 0.3$, coastalness scores were computed for 37 towns in Kent (treating Canterbury as a town rather than a city owing to population size). Fig. 4 shows the resulting coastal and inland classification. Table 5 reports the coastalness values for each town in Kent obtained using the proposed algorithm, alongside their corresponding binary coastal classification at the selected threshold.

5. Discussion

5.1. Summary of key findings and strengths of the approach

This study demonstrates that commonly used operationalisations of "coastal community" in public health research rely on spatial units and distance thresholds that do not map cleanly onto meaningful or coherent communities. In response, we developed and applied a population-weighted, continuous measure of coastalness defined at the level of bounded settlements rather than statistical output areas.

Applying this approach to towns in Kent shows that the proposed metric behaves in line with geographic intuition while retaining information lost by binary classifications. Towns conventionally regarded as coastal receive higher coastalness scores, inland towns cluster near zero, and intermediate cases are represented along a continuum rather than being forced into arbitrary categories. This allows distinctions to be made between strongly coastal, weakly coastal, and effectively inland settlements within the same analytical framework.

A key strength of the approach is that it explicitly incorporates settlement geometry and population distribution, rather than relying on single centroids or fixed distance bands. By weighting population exposure to the coast using a distance-decay function, the method captures internal heterogeneity within settlements and avoids sharp boundary effects introduced by binary thresholds.

Importantly, the proposed coastalness metric is transparent, reproducible, and adaptable. All assumptions, including boundary choice, coastline definition, and decay parameterisation are explicit and can be adjusted to suit different geographic contexts or research questions. While demonstrated here for towns in Kent, the method is readily transferable to other regions where appropriate settlement boundaries

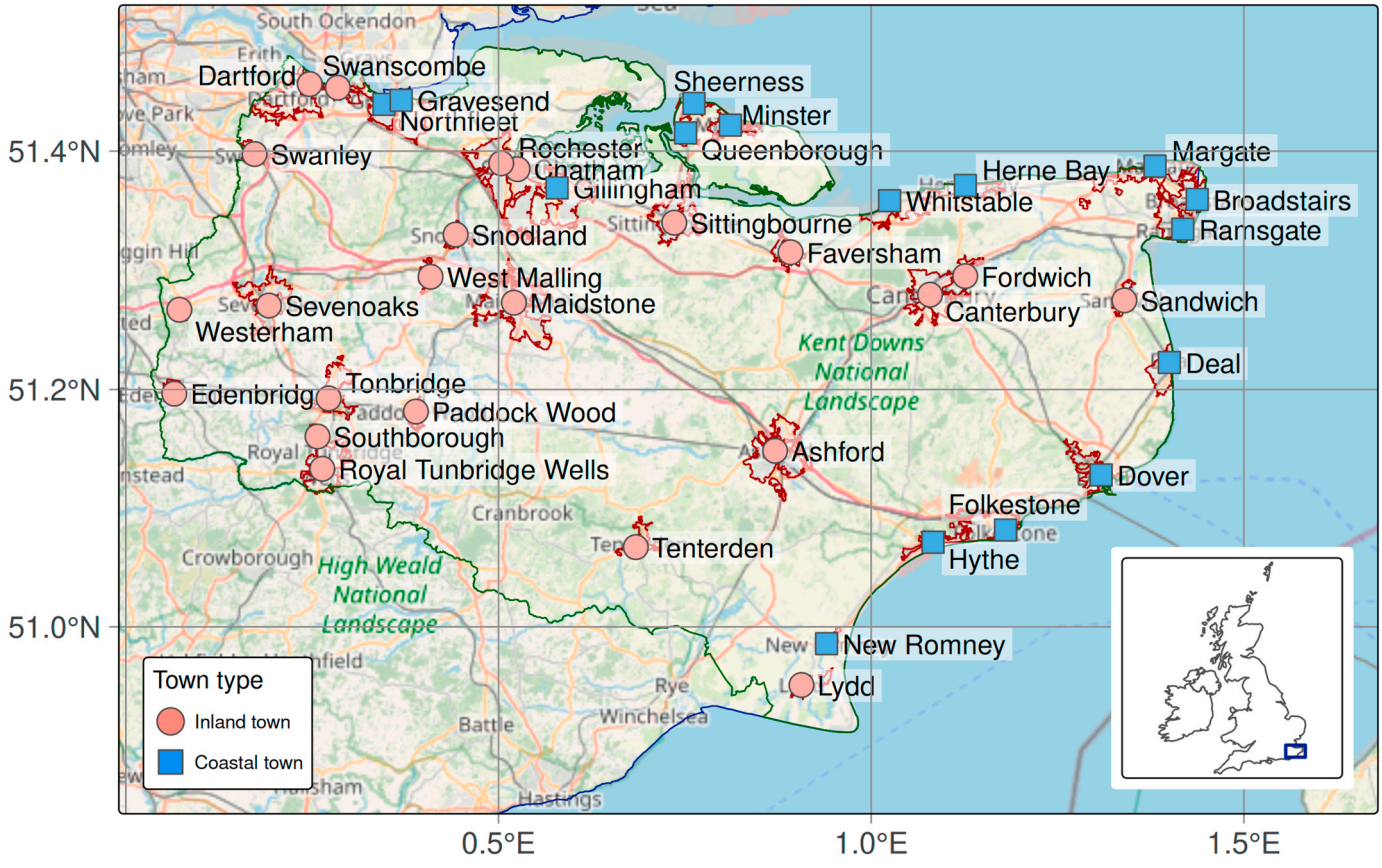

Fig. 4. Categorisation of towns in Kent into Coastal and Inland according to our algorithm.

Table 5
Town, coastalness, and coastal category at a threshold 0.5.

Town	Score	Class
Ashford	0.010	Inland
Broadstairs	0.735	Coastal
Canterbury	0.070	Inland
Chatham	0.338	Inland
Dartford	0.074	Inland
Deal	0.758	Coastal
Dover	0.612	Coastal
Edenbridge	3.9e-05	Inland
Faversham	0.412	Inland
Folkestone	0.740	Coastal
Fordwich	0.113	Inland
Gillingham	0.612	Coastal
Gravesend	0.641	Coastal
Herne Bay	0.869	Coastal
Hythe	0.808	Coastal
Lydd	0.329	Inland
Maidstone	0.018	Inland
Margate	0.813	Coastal
Minster	0.702	Coastal
New Romney	0.575	Coastal
Northfleet	0.643	Coastal
Paddock Wood	4.5e-04	Inland
Queenborough	0.924	Coastal
Ramsgate	0.728	Coastal
Rochester	0.475	Inland
Royal Tunbridge Wells	4.7e-05	Inland
Sandwich	0.391	Inland
Sevenoaks	0.001	Inland
Sheerness	0.856	Coastal
Sittingbourne	0.408	Inland
Snodland	0.066	Inland
Southborough	6.9e-05	Inland
Swanley	0.018	Inland
Swanscombe	0.354	Inland
Tenterden	0.006	Inland
Tonbridge	3.0e-04	Inland
West Malling	0.020	Inland
Westerham	2.6e-04	Inland
Whitstable	0.794	Coastal

and population data are available.

5.2. Interpretation, sensitivity, and boundary effects

To illustrate how the coastalness metric behaves in practice, we examine several towns that highlight its interpretability and sensitivity to geographic structure.

We can see, for example, that New Romney is classified as coastal at a threshold of 0.5, but its coastalness score (0.575) indicates a weak coastal association. This reflects the fact that a substantial proportion of the town's built-up area and population lies inland, with only limited proximity to the coast. The nearby coastal settlement Littlestone-on-Sea is often perceived as part of the wider local area, yet is classified as a separate village and are not included within the New Romney Built-Up Area. This illustrates how settlement boundary definitions directly shape measured coastal exposure, and how socially or geographically associated places may be excluded from boundary-based representations. The continuous coastalness measure captures this ambiguity in a way that would be obscured by a binary classification. This is illustrated in Fig. 5.

The coastalness metric is sensitive to how the coastline is delineated, particularly in estuarine environments. Decisions about where an estuary is considered to end and the open coast to begin therefore require contextual geographic judgement, as different cut-off choices can affect coastalness values for nearby settlements.

The metric is also sensitive to the internal spatial arrangement of residential areas within a settlement. Dover provides a useful example: despite being a major port and conventionally regarded as highly coastal, Dover has a relatively low coastalness score (0.612). The reason for this is illustrated in Fig. 6 which shows the distance from each LSOA population weighted centroid within the Dover BUA to the coast.

Dover's built-up area is narrow, with much of its residential population located inland. In addition, residential areas are physically separated from the coast by dock infrastructure and cliffs. The south-eastern portion of the BUA contains no population-weighted centroids, as it consists primarily of dock facilities, while the southern area is dominated by freight terminals.

This raises the question of whether port infrastructure or other physical barriers should be explicitly accounted for when calculating coastalness. One possible extension would be to incorporate additional weighting terms, for example: indicators of major port infrastructure or distance to functional coastal access points. We leave exploration of such refinements to future work.

5.3. Limitations

Several limitations of the proposed approach are inherent to the available geospatial data and modelling choices.

First, the analysis relies on Built-Up Areas (BUAs) as proxies for settlement boundaries. While BUAs provide spatially coherent representations of built development, they are derived from land-use classifications and design choices made by Ordnance Survey, including a minimum area threshold. As a result, smaller settlements, particularly hamlets, and fragmented or ribbon development may be underrepresented or excluded altogether. BUAs should therefore be understood as pragmatic boundary proxies rather than definitive representations of community extent.

Second, the method depends on population-weighted centroids to represent the spatial distribution of residents within a settlement. Although this approach improves on simple geometric centroids, it necessarily compresses heterogeneous residential patterns into a finite set of points, trading spatial fidelity for tractability and reproducibility. Settlements with elongated shapes, strong internal gradients, or physical barriers between residential areas and the coast may therefore have coastalness scores that underrepresent functional or symbolic coastal relationships.

Third, the approach inherits limitations from the settlement classification framework itself. Categories such as city, town, village, and hamlet are not defined using consistent or objective criteria, and are shaped by historical designation and administrative convention. While these labels are the only nationally standardised settlement categories available, they do not necessarily correspond to lived or socially meaningful notions of community in all cases.

Finally, the calculation of coastalness is sensitive to how the coastline is delineated, particularly in estuarine environments. Decisions about where an estuary is considered to end and the open coast to begin inevitably involve a degree of arbitrariness, and different cut-off choices may yield different coastalness values for nearby settlements. This sensitivity underscores the importance of contextual geographic knowledge when applying the method to new regions.

5.4. The role of human judgement

This work also raises broader questions about the role of human judgement in the construction of geospatial data and definitions.

Much of the extant geospatial data are derived from heuristics and local knowledge. For example, settlement centroids rely on *"the position that the majority of informed people would accept as being the 'centre' of the settlement"* (Ordnance Survey GB, 2023a) (p4). But how does ONS or any other organisation determine the consensus of the majority? Why did the ONS decide to use this definition: why not population centroid, address centroid, wealth centroid, or perhaps even road density?

Likewise, the ONS definition of BUAs depends on the National Land Use Database which itself is "validated" using "fieldwork" but the technical documentation leaves the details of this validation opaque.

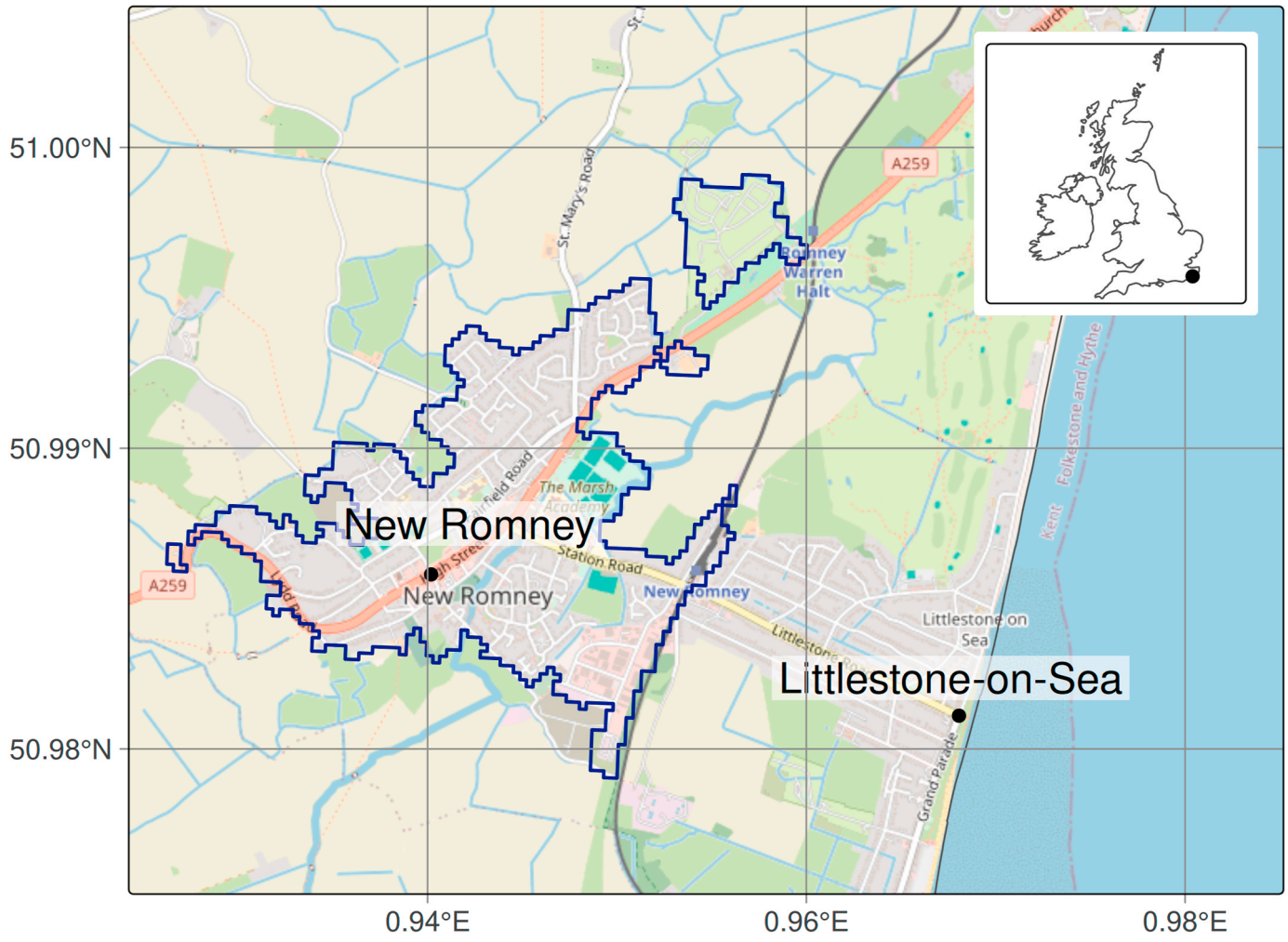

Fig. 5. Built-Up Area (BUA) boundary for New Romney (outlined) shown alongside the adjacent village of Littlestone-on-Sea, which is not assigned a BUA. Although Littlestone is often perceived as part of the wider coastal locality, its exclusion illustrates how settlement boundaries shape measured coastal exposure and contribute to weak coastalness scores despite geographic proximity.

In the work you are now reading, in lieu of "canonical" (ONS sanctioned) shoreline boundaries, we manually cut off large estuaries using our own local knowledge of when they stopped being coastal. Is this internal "fieldwork" any less valid than that of the ONS? What buys it that validity, authority alone? Or is it an assumption that the ONS must be better at this kind of thing because of its inherited expertise? These questions are important because we want to strip away bias and get to the heart of what it means to define something.

If we move beyond direct case-by-case human judgement we often find heuristics. The ONS heuristic differentiating coastal and non-coastal BUAs described above is declared without explanation: it stands only in the strength of its own internal coherence. The heuristic itself is grounded in human reasoning, validated against *a priori* knowledge of what should be coastal and what shouldn't be, and then adjusted to fit.

Is this any more objective than explicitly deciding coastalness on a case-by-case basis? If not, what do we gain from heuristics: consistency and a "defensible" story that *looks* more methodological? And what do we lose: specificity, precision?

If we really wanted to be "objective", how would we approach this? There are no axioms buried in the soil for us to find, no boundary lines etched into the bedrock denoting truth. Centroids, boundaries, and categories are not found: they are made. Even seemingly un-grounded fields such as mathematics reveal this: Euclid defined his postulates and flat geometry followed, but contest the postulates and Riemannian space emerges. All knowledge is grounded in human axiomatic decisions.

Once we admit this we can understand that there is no "true" answer to the question "what is a coastal town?". Instead we can fall back on the question: what purpose does a definition serve?

In the context of public health, our definitions serve utility. Our objective is to improve public health and distinctions between entities matter only in so much as they serve this goal.

The hypothesis underlying the concern around coastal communities stems from the idea that somehow coastalness *matters*, that it is *causal* in some fundamental way or at least a proxy for other presumed causal factors such as access to resources.

It would be feasible to manually label every town in the UK as either coastal or non-coastal simply by traversing the coast and using local knowledge in each case. For a professional organisation such as ONS, this could be done fairly quickly.

A problem with this approach is that human judgement can confound causal variables, and reasoning is hidden unless we ask for it to be expressed. If a human judges two towns that are roughly the same distance from the coast but considers one coastal and another non-coastal, what is the deciding factor?

This is both the strength and weakness of human intuition: intuition encodes latent meaning and can be very powerful, but is not explanatory by itself and can be biased. We can however ask a human what makes them differentiate similar instances, and in the best case, these distinctions can be made explicit and formalised into measurable criteria.

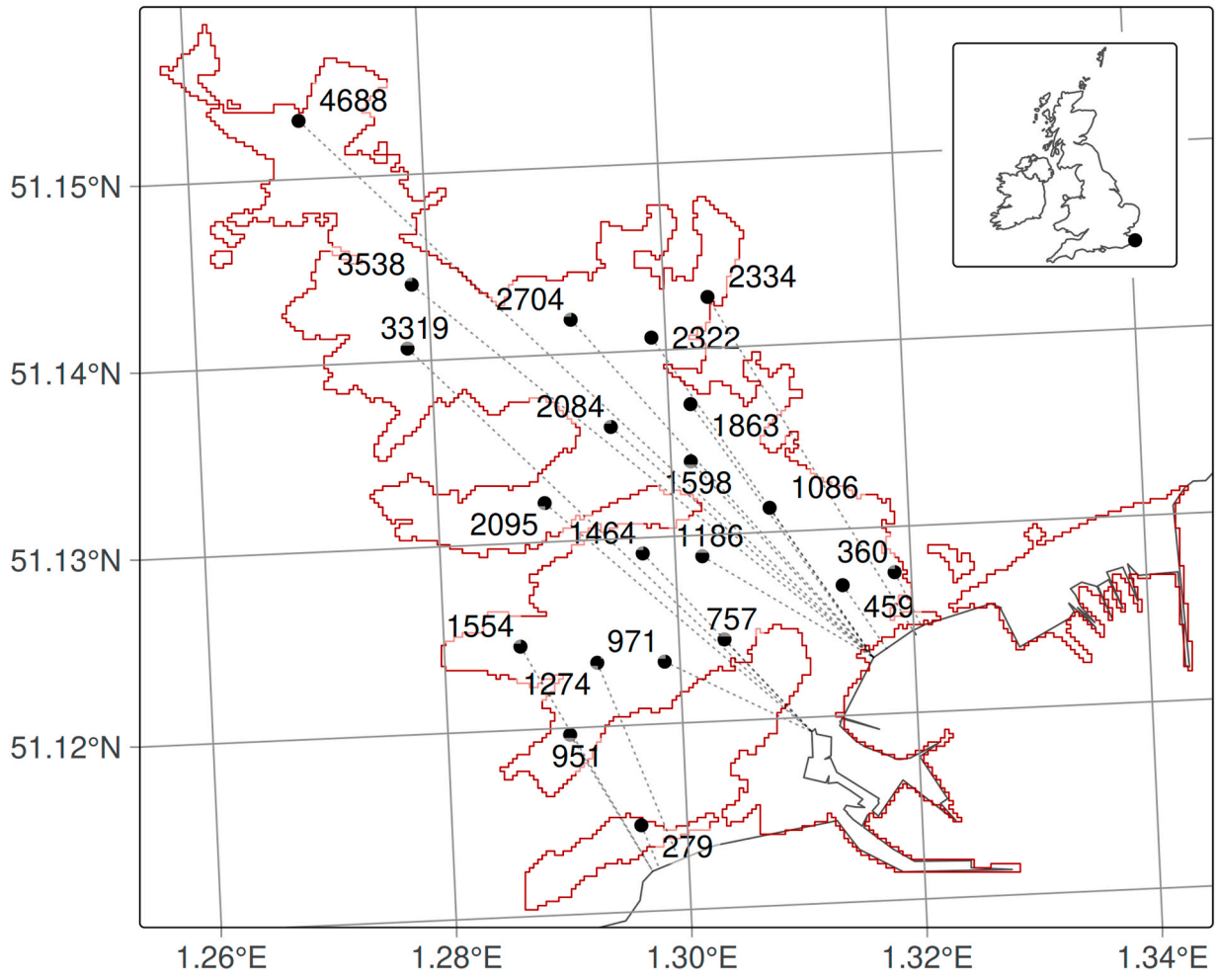

Fig. 6. Distance in metres for population weighted centroids of Dover BUA to the coast. Plot uses 2024 BUAs and 2021 LSOA population weighted centroids (latest release as time of analysis).

Our goal should be to remove ambiguity and reduce our definitions to what can be physically measured. The creation of discrete categories "coastal" and "non-coastal" is analogous to intuition: it might encode something important but every abstraction introduces compromise.

In the worst case, intuition can completely mislead. Imagine someone who thinks that the average horoscope of a town is important to health and includes it in a model. There is no evidence to suggest that this would have a meaningful impact on health, but in some reasoning systems, horoscopes matter. Including such a variable not only risks detecting false associations in small samples, but any natural imbalances will take power away from the actual causal variables as the model accounts for the variance.

We seek parsimony in modeling, not because of aesthetics, but because each extraneous non-causal variable adds noise and weakens the signal of the actual determinants.

If we follow this argument through, what does it mean for definitions? It means we should not conflate geospatial constructs with human intuition unless there is no better way.

This is why we have presented a geospatial definition of coastalness which is a continuous variable, so that the geospatial coastalness of a place can be isolated from other factors and the relative causal weight of covariates for a given research outcome disentangled.

In the Kent example presented above, we apply a binary classification on top of the continuous coastalness metric. This is not because we consider binary thresholds to be epistemically preferable, but because such classifications remain unavoidable in policy and applied research contexts, where decisions are typically framed in terms of "coastal" versus "inland" areas. The use of *a priori* geographic knowledge to inform the selection of a threshold does introduce a degree of circularity, and we acknowledge this explicitly. However, this circularity is not unique to our approach; it is inherent in any threshold-based definition of coastal status, including those currently used in policy and official statistics.

Crucially, the continuous coastalness metric allows analyses to proceed without reliance on this threshold at all. Where binary classification is required, thresholding can be treated as a pragmatic reduction rather than a claim about natural categories. In principle, threshold selection could also be informed empirically with respect to a specific outcome, for example by choosing a value of λ that optimises discrimination or maximises information gain, rather than relying on intuition alone. We retain the continuous formulation precisely to preserve this flexibility and to avoid conflating methodological convenience with conceptual truth.

We also used human knowledge to bound the influence of estuaries, and we depended on heuristically derived concepts: settlement centroid and BUA boundary. We should strive to identify such biases and work towards eliminating them. We hope that this paper serves as a grounding for further objectification.

5.5. *Boundaries*

Beyond the specific coastalness metric, our findings expose a more fundamental problem: the absence of canonical settlement boundaries in the UK.

The lack of "canonical" settlement boundaries is not just a technical limitation: it is a foundational conceptual problem. Without clear boundaries for settlements we cannot measure their characteristics, analyse differences between settlements, or target treatments towards ameliorating any differences that are observed. We used BUAs as substitutes for settlement boundaries. In this section we will discuss the limitations of this approach and the way forward.

We have shown that current BUA membership is not exclusive: places can share BUAs, Fig. 7 shows Kingswood and Fishponds BUA which contains 11 individual villages (for reference, this BUA contains no towns).

We have shown that BUAs are however reasonable proxies for the populations of Towns; 87.66% having a one-to-one mapping between town and BUA and 99.83% having a one-to-one mapping if we allow villages and settlements to be subsumed by them. Therefore, we argue that our approach is suitable for aggregating data and analysing differences at the level of Towns, including through the lens of coastalness as a covariate.

BUAs are less suitable for villages and hamlets, due to the relative lack of one-to-one mappings, but for particular analyses may be sufficient if the requisite power is met.

Whilst we allowed town BUAs to subsume villages in our definition of coastalness for towns in Kent, further research is needed to understand exactly what this means for the semantics of place.

Subsumption does not preclude the need for more granular boundaries. This is especially true for cities and very large towns where regions and boroughs are often larger than individual towns and can have distinct populations and needs. For example: we may want to compare places within Greater London, and also analyse Greater London as a whole. This raises questions of how to subdivide BUAs and how to aggregate them.

Looking at the BUA boundary data, there is considerable variation in how different places are treated. Birmingham appears to have one main BUA encompassing most of its urban population, whereas Bristol appears to be split into two main BUAs, and London has at least 25 BUAs.

When we seek explanation for these differences we can find no epistemically grounded justification, and have to turn instead to the heuristics underpinning the construction of BUAs themselves, the presumed authority of the NLUD data, and the artifacts that emerge when the (somewhat latent) rules are applied. A visual inspection of BUAs, reveals inconsistencies, for example in some cases a BUA will be split on an A road boundary, yet in other cases areas are split on smaller B roads despite being bisected by A roads.

A final point is that BUA boundaries do not necessarily align with administrative boundaries; a single BUA can fall into more than one administrative area. Since it is administrations that are expected to take action and be responsible for serving an area, future work should examine the relationship between the definition of settlement extents and the functional relationships they have with abstract servicing entities such as councils and NHS trusts.

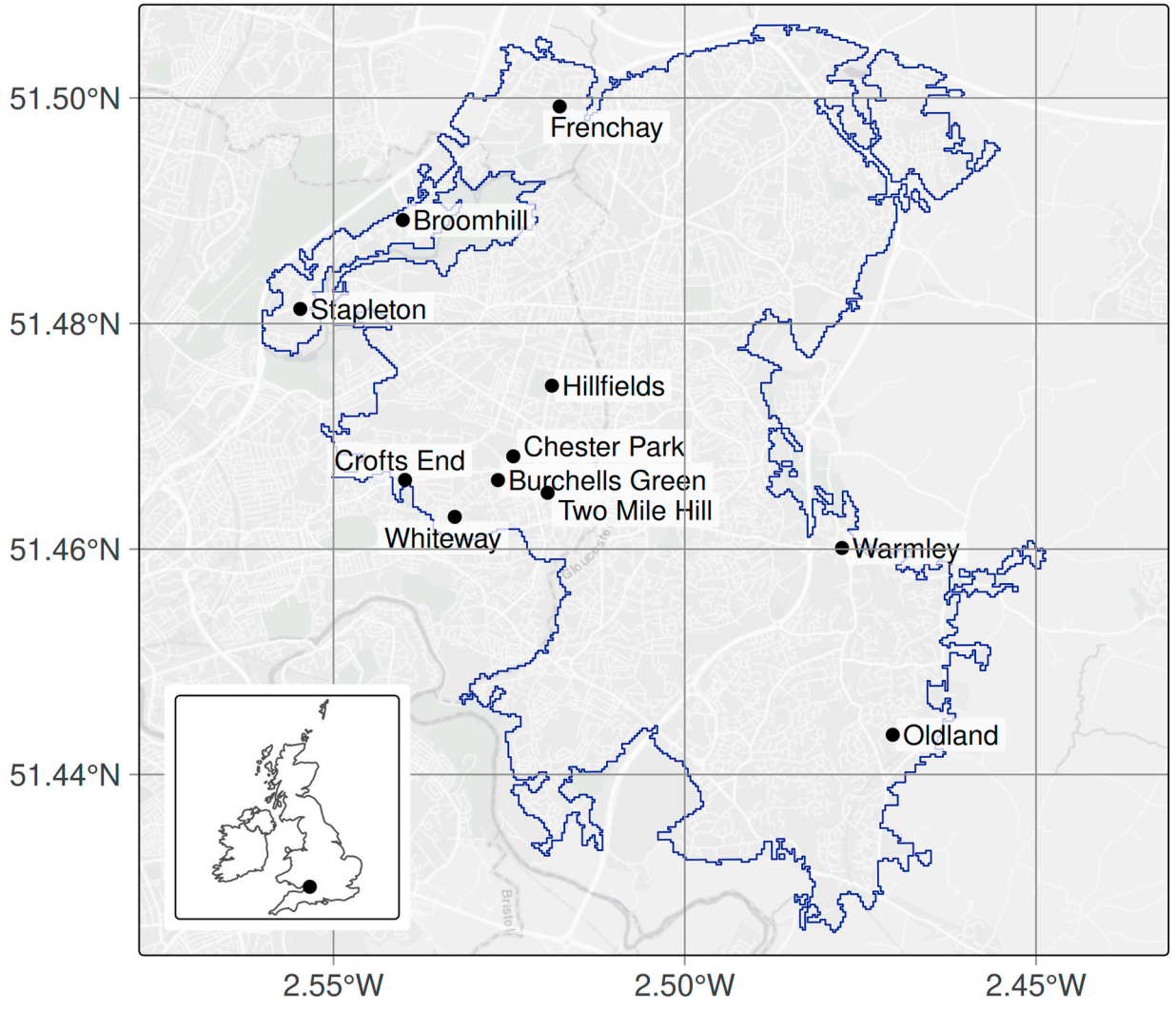

Fig. 7. Kingswood and Fishponds BUA containing 11 individual villages.

All of this points toward the urgent need for better settlement boundaries. We need boundaries for places at every scale: from the smallest hamlet to the largest city, and we need clear rules for subsumption and hierarchical membership tied to the semantics of place.

This is something that would naturally fall within the domain of Ordnance Survey, and we encourage the UK government to fund the provision of essential data for settlement-level epidemiology in this regard.

5.6. Places

The ONS provides settlement categories (city, town, village, hamlet, other) and we used them as our canonical reference. However, these categories are not consistently grounded in objective criteria. City status, for example, derives from historical designation rather than size; towns are defined by thresholds of contiguous built-up area irrespective of population; and villages and hamlets are only defined in relation to each other ("smaller than towns," "smaller than villages"), without quantitative thresholds.

Given this, it is more coherent to treat settlements directly as spatial extents with associated populations, rather than relying on categorical definitions that are arbitrary and inconsistently applied.

Treating settlements as extents with populations makes sense, but since ONS settlement points lack boundaries, population estimation requires constructing those extents first. Hence the problems of boundary definition and population assignment are inseparable and need to be addressed together in future work.

5.7. Coastalness

When we looked at the coastalness of the towns in Kent we observed that Dover was weakly coastal having a coastalness of 0.612 where the "is coastal" threshold 0.5 was preconditioned on our *a priori* understanding. Given that Dover is a well known international harbour and ferry port we might expect it to be strongly coastal, but it is precisely this harbour status that reduces its coastalness, because residential areas are structurally restricted from the coast by the harbour itself.

We suggested that the coastalness algorithm could be adjusted with a weighting term to account for harbours, but the observation reveals something more fundamental. In the case of Dover we can think of the harbour as being an extension of the coast; an extension of the structural and functional exclusion that coast induces in relation to a place.

It is not unreasonable to argue that the coast might impose structural limits on the growth and connectivity of a place, and that this might contribute to factors related to overall health, such as access to opportunity or healthcare. But whether coastal barriers are unique in this regard is a different question.

It may be the case that the geospatial constraints the coast imposes generalise to other structural barriers: mountain ranges, national parks, airports, military institutions, industrial estates, etc. In which case we might decide that not all inland areas should be treated equally and it may be that we are better to characterise every place by a more abstract concept of structural constraints. This opens another direction for future research.

6. Conclusion

We have presented an overview of the issues that arise when trying to define the coast, the concept of community, and their combination in coastal community. We have presented a geospatial definition for community, bounded communities using existing public boundary data, and presented an algorithm for computing a continuous measure of coastalness for bounded areas. We demonstrated the effectiveness and flexibility of this method for Towns in Kent. We have argued for the construction of comprehensive canonical settlement boundaries to enable future analytical work to be grounded on epistemically solid foundations. Though grounded in the specific case of Kent, the methods and reflections we offer are extensible across spatial contexts, and invite a broader rethinking of how geospatial knowledge is constructed and applied.

CRediT authorship contribution statement

Ashley Mills: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Visualization. **Salmaan Ansari:** Conceptualization, Methodology, Writing – review & editing. **Maja Niksic:** Conceptualization, Writing – review & editing. **Lindsay Forbes:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition.

Ethical statement

This study is a desk-based methodological and conceptual analysis using publicly available, aggregated geospatial and population data. No individual-level, identifiable, or sensitive personal data were used. As such, the study does not involve human participants and did not require review or approval by a research ethics committee.

Declaration of the use of AI

An AI-assisted tool (ChatGPT, OpenAI) was used during manuscript preparation to support clarity of expression. Any AI-generated suggestions were critically reviewed and, where appropriate, adapted by the authors, who take full responsibility for the manuscript. The conceptual development, analysis, and substantive arguments originated from the authors.

Funding

This research was supported by the National Institute for Health and Care Research (NIHR) Applied Research Collaboration Kent, Surrey, Sussex.

The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Ashley Mills reports financial support was provided by National Institute for Health and Care Research (NIHR) Applied Research Collaboration Kent, Surrey, Sussex. Lindsay Forbes reports financial support was provided by National Institute for Health and Care Research (NIHR) Applied Research Collaboration Kent, Surrey, Sussex. Salmaan Ansari reports financial support was provided by National Institute for Health and Care Research (NIHR) Applied Research Collaboration Kent, Surrey, Sussex. Maja Niksić reports financial support was provided by National Institute for Health and Care Research (NIHR) Applied Research Collaboration Kent, Surrey, Sussex. If there are other authors, they 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

This research was funded by the National Institute for Health and Care Research (NIHR) Applied Research Collaboration Kent, Surrey, Sussex.

The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Data availability

All data used in this manuscript is publicly available and is linked to within the manuscript.

References

Asthana, S., & Gibson, A. (2022). Averting a public health crisis in England's coastal communities: A call for public health research and policy. *Journal of Public Health, 44*(3), 642–650. https://doi.org/10.1093/pubmed/fdab130

DHSC. (2021). *Chief medical Officer's annual report 2021: Health in coastal communities*. https://www.gov.uk/government/publications/chief-medical-officers-annual-report-2021-health-in-coastal-communities.

Gauvain, H. (1935). Sea Bathing in the treatment of surgical tuberculosis. *British Medical Journal, 2*(3909), 1087–1090. https://doi.org/10.1136/bmj.2.3909.1087

Geiger, S. J., White, M. P., Davison, S. M. C., Zhang, L., McMeel, O., Kellett, P., & Fleming, L. E. (2023). Coastal proximity and visits are associated with better health but May not buffer health inequalities. *Communications Earth & Environment, 4*(1), 1–9. https://doi.org/10.1038/s43247-023-00818-1

Hunt, E. (2019). Blue spaces: Why time spent near water is the secret of happiness. *The Guardian*. https://www.theguardian.com/lifeandstyle/2019/nov/03/blue-space-living-near-water-good-secret-of-happiness.

Jewkes, R., & Murcott, A. (1996). Meanings of community. *Social Science & Medicine, 43*(4), 555–563. https://doi.org/10.1016/0277-9536(95)00439-4

MacQueen, K. M., McLellan, E., Metzger, D. S., Kegeles, S., Strauss, R. P., Scotti, R., Blanchard, L., & Trotter, R. T. (2001). What is community? An evidence-based definition for participatory public health. *American Journal of Public Health, 91*(12), 1929–1938. https://doi.org/10.2105/AJPH.91.12.1929

Office for National Statistics. (2023). A Beginner's guide to UK geography: Geographical areas frequently used in the publication of official statistics. https://open-geography-portalx-ons.hub.arcgis.com/datasets/ons::a-beginners-guide-to-uk-geography-2023-1.

Office of the Deputy Prime Minister. (2006). National land use database: Land use and land cover classification. https://assets.publishing.service.gov.uk/media/5a78b635ed915d07d35b1d52/144275.pdf.

O'Grady, S. (2016). REVEALED: The mental health benefits of living by the seaside. https://www.express.co.uk/life-style/property/665761/Revealed-mental-health-benefits-living-seaside-house-sea-life.

ONS. (2024). Coastal communities, characteristics of built-up areas, England and Wales: Census 2021. https://www.ons.gov.uk/peoplepopulationandcommunity/housing/articles/coastalcommunitiesandcharacteristicsofbuiltupareasenglandandwales/census2021.

ONS. (2025a). ONS digital boundaries. https://www.ons.gov.uk/methodology/geography/geographicalproducts/digitalboundaries.

ONS. (2025b). ONS GB shoreline and related work - GitHub repository. https://github.com/ONSgeo/shoreline.

ONS. (2025c). *Open Geography Portal*. https://geoportal.statistics.gov.uk/.

Ordnance Survey. (2022). OS open built up AREAS™ – technical specification. https://www.ordnancesurvey.co.uk/documents/product-support/tech-spec/os-open-built-up-areas-technical-specification.pdf.

Ordnance Survey GB. (2023a). OS open NAMES™ – technical specification. https://www.ordnancesurvey.co.uk/documents/product-support/tech-spec/os-open-names-technical-specification-v2.2.pdf.

Ordnance Survey GB. (2023b). OS open names. https://osdatahub.os.uk/downloads/open/OpenNames.

Valen, D. (2022). Toward an architecture of air and steam: Climate, bathing, and public health in victorian britain. *Journal of Architectural Education, 76*(1), 21–36. https://doi.org/10.1080/10464883.2022.2017688

Welsh, C., Welham, C., Anderson, J., Green, M. A., Quinn, C., Lai, J., Vernon, S., & Paley, L. (2024). Can we empirically derive a geographic definition of “coastal” for use in cancer data reporting? An ecological modelling study using England's national cancer registry. *BMJ Public Health, 2*(2). https://doi.org/10.1136/bmjph-2024-001067

Wheeler, B. W., White, M., Stahl-Timmins, W., & Depledge, M. H. (2012). Does living by the Coast improve health and wellbeing? *Health & Place, 18*(5), 1198–1201. https://doi.org/10.1016/j.healthplace.2012.06.015