Street Art & Urban Creativity | Vol. 12, No. 7, 2026 | ISSN 2183-9956
Street Art & Urban Creativity. Scientific Journal

AI-Enabled Remote Sensing for Smart Cities Covering from Hyperlocal to the HyperSky

Evidence and opportunities across urban heat island, car parks, environmental monitoring and urban resilience for wildfires

Iris Cuevas Martinez[1], Antonio J. Jara[1]
[1]Libelium Comunicaciones Distribuidas S.L., Zaragoza, 50018, Spain
Correspondence: i.cuevas@libelium.com

KEYWORDS

Earth observation
Urban planning
Urban heat islands
Car-park monitoring
Wildfire intelligence
Internet of Things
Very low Earth orbit

ABSTRACT

Smart cities need observations that connect regional change with local intervention. Remote sensing can combine satellite coverage, drone detail and hyperlocal sensors (Internet of Things) to improve urban planning. This evidence-based synthesis examines parking, urban heat, wildfire and air-quality applications, separating validated results from commercial specifications and scenarios. European thermal missions and very low Earth orbit (VLEO) initiatives expand opportunities for urban planning. A decision-oriented architecture links these capabilities to governed, quality-qualified services without presenting simulations or provider specifications as field validation. We propose decision-specific validation and adaptive acquisition, identifying VLEO opportunities while distinguishing emerging capabilities from operational services. Value depends on usable coverage, uncertainty and timely urban action.

1. Introduction

Urban planning requires observations that connect individual places with the wider systems shaping them. A parking space is part of a mobility network and a land allocation decision; a hot pavement belongs to a neighbourhood heat pattern; a fire near the urban fringe can interrupt transport, damage utilities and expose distant residents to smoke. Cities therefore need both local measurements and a consistent view of their territory. The useful opportunity in remote sensing is to connect these scales so that a city can identify where intervention matters, investigate the selected locations and measure whether the intervention works. An urban authority rarely needs an image for its own sake. It needs evidence to decide where to add shade, whether a car park is overprovided, which access routes remain usable during a wildfire, or where environmental monitoring is missing. These questions extend beyond the footprints of existing ground installations. The central opportunity for Earth observation (EO) is therefore not to replace every local sensor, but to make otherwise unobserved urban territory part of a consistent planning evidence base. The documented BeatTheHeat deployment demonstrates this integration for urban heat, while published parking and wildfire studies provide complementary evidence of spatial scalability and task-specific accuracy (BeatTheHeat, Cuevas et al., 2026; Drouyer & de Franchis, 2020; Xu et al., 2026).

This article calls that connection the hyperlocal-to-hypersky continuum. Hyperlocal sensing includes fixed cameras, parking sensors, environmental stations and field instruments. Drones provide flexible observations of sites and corridors. Earth observation (EO) adds repeated views of municipalities, metropolitan regions and their hinterlands. Very low Earth orbit (VLEO) introduces the prospect of finer orbital observation, while conventional low Earth orbit (LEO) and geostationary systems remain essential for wide-area context and surveillance. Hypersky is used here as an organising concept for this combined observation portfolio, not as a new spectral band or a standard orbital classification. The next Figure introduces this complementarity of sensors capabilities.

Figure 1. The hyperlocal-to-hypersky continuum remote sensing capabilities



Artificial intelligence (AI) makes the continuum useful by converting heterogeneous observations into indicators: vehicle detections, parking occupancy, surface classifications, temperature estimates, smoke or fire masks, and spatial predictions of pollution. However, a

detailed image is not itself a reliable planning product. The target variable, reference measurement, spatial support and decision deadline must be defined before choosing the sensor or model. A municipality purchasing a heat-risk service needs to know whether the map represents land surface temperature or human exposure; a mobility department needs to know whether a parking indicator is a snapshot or a continuously updated service.

The distinction is demonstrated by published work. Satellite vehicle detection has been applied across Barcelona, satellite parking occupancy has been evaluated in Alkmaar, and drone and satellite temperatures have been compared with field measurements in an urban park in Sejong. These studies establish practical opportunities while revealing different limitations of spatial resolution, canopy, shadow and sampling (Sheehan et al., 2023; Stopic et al., 2023; Kim et al., 2021). Their results should inform deployment choices without being converted into an unsupported league table of technologies.

The contribution is a decision-oriented synthesis with reproducible secondary calculations. It extends mobility observability to urban heat, wildfire exposure and air quality, and connects these applications to a common acquisition and validation framework. Three questions guide the analysis. Which urban decisions benefit from each observation scale? What do real evaluations establish about accuracy and scalability, and which comparisons remain invalid? How can cities and EO service providers build useful new services around emerging orbital capabilities while retaining reliable local evidence?

The central proposition is that a scalable urban service uses broad coverage to allocate attention and uses local detail to resolve consequential uncertainty. This does not prescribe the same sensor sequence for every problem. A slowly changing inventory can rely mainly on orbital observations, whereas rapidly changing parking guidance or emergency response may require continuous local sensing. The appropriate portfolio follows the decision and its consequences.

2. Materials, evidence and comparative method

2.1. Methodology

The research design combines analysis of an operational case with a targeted comparison of published evaluations and current technical documentation. This work is a structured applied synthesis of the evidences, opportunities, use-cases and how Smart Cities will take benefit of new emerging technologies beyond the IoT for sensing, as satellite and drones; which aims to inspire, analyse and identify the emerging opportunities to address urban planning and local digital twins with new data sources, that before were not imaginable. Today, satellite capabilities, drones' capabilities and the IoT sensors evolution are enabling new paradigms and solutions that thanks to the AI evolution are creating high scale capabilities. Face to the 2030 we can really think about world-scale databases as the EU LDT toolbox European building database, the EU Copernicus DataSpace for remote sensing and satellite-based large datasets, enabling even near-real time monitoring as Meteosat every 10 minutes updates of satellite information, and that together with mission critical systems as drones, are able to monitor any location of the planer in few hours. It is also highly remarkable the evolution of VLEO and LEO missions with a lot of emerging companies with satellite technologies that offer 6h to 24h monitoring capabilities with accuracy under 35cms pe pixel, what makes really feasible, thanks to AI, to identify green zones, car parks, wildfires, people location, floods effects, buildings status, and much more thanks to AI and computer vision capacity.

Table 1 defines the evidence classes used throughout. Reported case measurements are retained with their original limitations.

Table 1. Evidence classes and permitted interpretation

Class	Evidence used	Interpretation boundary
Operational case	Implementation and partner-component results	Demonstrates deployed integration; not independent validation of every layer.
Published evaluation	Satellite, UAV (drones), and IoT studies with numerical results	Supports the reported task and test setting; no pooled cross-study ranking.
Official specification	Mission, provider and collection documentation	Describes sampling, access and roadmap; does not prove decision accuracy.
Analytical scenario	Coverage arithmetic and value-of-information reasoning	Explains a design mechanism; is not a measured cost saving or service level.

Each result is extracted against five descriptors: the measurand, spatial support, temporal support, reference procedure and operational extent. The measurand is what is estimated, such as surface temperature rather than air temperature. Spatial support is the actual area represented by a measurement, not simply the displayed raster cell. Temporal support includes acquisition time, averaging interval and reporting delay. Reference quality covers calibration, annotation, synchronization and whether evaluation data were held out. Operational extent distinguishes a validated site from a larger rollout that lacks equivalent ground truth.

This procedure separates three forms of comparability. A paired experiment can compare methods within its stated conditions. A cross-study comparison can describe different demonstrated capabilities but cannot attribute all performance differences to platform choice. A service comparison asks whether the result arrives over enough of the required area, at acceptable cost and age, to support a specified decision. The paper uses the first where a paired thermal study exists, the second for heterogeneous parking and wildfire evaluations, and the third to develop the common architecture.

Targeted searches covered satellite parking occupancy, city-scale vehicle detection, UAV parking monitoring, urban thermal validation, active-fire segmentation, ground sensor calibration and satellite-assisted pollution estimation. Studies were selected when they provided an identifiable application, a stated reference or test set, and a quantitative result relevant to an urban decision. Mission documentation was used to describe measurement capabilities. This purposive selection supports comparison of representative evidence.

Four evidence classes are kept distinct throughout the article. Published empirical results describe measurements or model evaluations reported by the cited investigators. Secondary calculations are deterministic transformations of their published numbers. Operational specifications describe documented sensor or product characteristics, which do not establish application accuracy. Design scenarios state assumptions used to explore a prospective service.

For each empirical comparison, we record the task, study setting, spatial or temporal support, reference, metric and principal limitation. This is essential because occupancy classification accuracy, object-detection F1, pixel-level fire precision and temperature RMSE measure different things. A high value on one task cannot be used to rank a modality on another. Table 2 summarises the observation roles before the application-specific evidence is examined.

Table 2. Observation roles across the urban sensing continuum

Observation layer	Spatial and temporal role	Urban contribution and main limit
Hyperlocal sensors and cameras	Persistent observation of installed points or views	Local state and reference measurements; coverage depends on deployment and representativeness.
Drone optical and thermal	Fine site or corridor surveys during permitted flights	Materials, geometry and targeted inspection; flight access and operating effort constrain repetition.
LEO optical imagery	Common snapshots across sites and urban networks	Inventory and visible objects; cloud, shadow and acquisition timing limit observability.
LEO thermal and SAR	Modality-specific regional or targeted observations	Thermal patterns or microwave scattering; native support and physical variables differ.
Geostationary observations	Repeated broad-area surveillance	Evolving regional context; coarse spatial support limits urban detail.
Emerging VLEO imagery	Potentially finer orbital detail	Small urban features; application accuracy and routine service availability require validation.

2.2. Making accuracy and scalability comparable

Accuracy must refer to an observable and a reference at compatible scales. Ground sampling distance (GSD) describes pixel spacing; it is not positional accuracy, temperature accuracy or classification performance. Native measurement support must also be distinguished from a resampled or downscaled output grid. Landsat 8 TIRS thermal bands have 100 m sampling, although thermal products can be distributed on a finer grid; its multispectral and thermal bands have different native resolutions.

For parking, the comparison unit can be a vehicle, a mapped space, a lot-level count or an occupancy fraction. Let O be occupied observable spaces, C be valid mapped capacity and U be spaces whose state cannot be observed. Observable occupancy is $O/(C - U)$, and the observable fraction is $(C - U)/C$. Both quantities should be reported. A model that excludes tree-covered spaces can improve its classification score while reducing coverage. Its citywide usefulness depends on both changes, not the score alone.

For temperature, remote measurements should first be compared with surface measurements representing the same material and time. A calibrated air-temperature sensor measures a different physical quantity and is not a direct reference for a roof or pavement temperature. To compare drone and satellite rasters, the fine observations should be aggregated to the satellite footprint or point-spread support before estimating bias and error. Retaining the fine image is still useful for examining variation within that footprint. Comparing a point on a wooden deck with a mixed pixel dominated by grass otherwise confounds retrieval error with representativeness.

Scalability includes usable area, observation frequency, processing, staff effort and delivery. We define usable area-time coverage as the sum, across planned observations, of area that passes quality and deadline requirements divided by total requested area across those observations. It is separate from the conditional accuracy achieved where measurements are available. A complementary service indicator is fully loaded cost divided by the number of accepted, timely observations of the specified decision unit. These definitions prevent a broad but intermittently unusable product and a narrow continuous sensor from appearing equivalent.

The age of information is the decision time minus acquisition time. It must not be replaced by revisit opportunity or processing speed. Delivery time includes acquisition waiting, transmission, processing, quality control and integration. For repeated images, an acquisition schedule can also introduce systematic bias: observing only clear mornings may miss peak afternoon heat, evening parking demand or short-lived smoke exposure.

3. From observation scales to urban decisions

3.1. Hyperlocal continuity and drone investigation

Hyperlocal systems provide temporal continuity where observations have direct operational value. A parking sensor can report a change at a designated space; a fixed camera can observe a group of spaces or a road approach; an environmental station can sample local conditions across the day. Their limiting factor is spatial representativeness. A dense time series at one point does not describe every street, and a citywide network requires installation, communications, maintenance and periodic calibration. These systems are especially valuable as reference sites and as persistent monitors at places where errors matter most.

Drones occupy a complementary role. They can inspect a shortlisted car park, thermal anomaly, drainage corridor or damaged asset with flexible geometry and high spatial detail. A thermal survey can resolve individual materials; oblique imagery can reveal features concealed from an orbital view. However, the observation remains dependent on flight windows, crew or autonomous operating arrangements, battery endurance, local obstacles and permissions. In Europe, operating conditions depend on the applicable drone category and geographical restrictions (European Union Aviation Safety Agency). A fast inference model does not remove these acquisition constraints.

The implication for deployment is to reserve repeated local observation for locations with high decision value, while using mobile surveys to answer targeted questions. For example, a city might install persistent sensors at heat refuges and public-transport interchanges, then use drones to investigate the surrounding surface conditions. The camera, field instrument and drone should share geospatial identifiers and timestamps. Otherwise, their nominally complementary outputs may be impossible to compare when a planning team needs to explain a discrepancy.

3.2. Orbital coverage and the VLEO opportunity

Orbital imagery supports consistent observation across administrative boundaries and across sites without installed instrumentation. Its archive can reveal land-use change and infrastructure development, while new acquisitions can describe a common spatial snapshot. Optical, thermal and synthetic-aperture-radar (SAR) observations answer different questions. Optical imagery supports geometry and object recognition; thermal observations describe emitted radiation and related thermal properties; SAR measures microwave scattering and supports weather-resilient structural or surface-change analysis. None of these modalities is a universal substitute for the others.

VLEO can increase ground detail for a given optical design by reducing the distance to the target. The prospect is attractive for narrow urban features, small assets, canopy gaps and parking geometry. Yet finer optical sampling does not automatically provide equivalent thermal detail, reliable small-object recognition, continuous access or better geolocation. Orbital drag, propulsion, atomic oxygen exposure, pointing stability, downlink and constellation size remain part of the service design. Lower altitude is one engineering parameter in an end-to-end observation system.

The commercial evidence illustrates this distinction. EOI Space describes 15 cm imagery for its Stingray programme; it is treated here as a provider-stated capability, not a validated city service (EOI Space). Albedo's Clarity-1 retrospective reports successful VLEO operations and thermal imaging, but explains that imaging problems prevented achievement of the complete 10 cm visible-imagery goal and that contact was subsequently lost. A flight-proven bus therefore does not establish a routine 10 cm municipal data supply (Haddad, 2026).

The geometric opportunity remains substantial. For an assumed 4.5 m by 1.8 m car, nominal support increases from 9 by 3.6 pixels at 0.50 m GSD to 30 by 12 pixels at 0.15 m. This is approximately 11.1 times the pixel area. At 0.10 m it becomes 45 by 18 pixels. These calculated footprints indicate additional information potentially available to a detector; they are not

predicted F1 values. Object contrast, motion, shadows, image processing and annotation still determine recognition quality. VLEO should consequently be evaluated through delivered native products and local validation, while its emerging applications are developed in parallel.

3.3. AI products that retain their measurement meaning

A useful processing chain preserves the distinction between observation and inference. For parking, georeferenced imagery is tiled at a scale that retains vehicles, detections are reconciled across overlaps, and objects are intersected with mapped capacity. For heat, radiometric quality and surface context precede aggregation or predictive downscaling. For wildfire, a detected anomaly is associated with time, location, fuel, weather and exposed assets before any response recommendation. For pollution, remote atmospheric variables are combined with meteorology and independent ground measurements to estimate surface concentrations.

A local digital twin can then link these indicators to planning objects, such as a school, a street segment or a park polygon. It should retain acquisition time, uncertainty, missing-data masks and the processing version alongside the indicator. A prediction can fill a service gap if it is explicitly marked as a prediction and its uncertainty is acceptable. It should not overwrite the distinction between a fresh observation and an estimated current state. This is particularly relevant to interactive AI interfaces, whose fluent descriptions can otherwise conceal weak or outdated evidence.

3.4. Resolution, coverage and freshness are separate properties

Spatial detail is valuable only in relation to a task. Detecting vehicles in an open car park requires different information from mapping thermal patterns or recognizing a fire-related anomaly. Ground sampling distance does not specify geolocation accuracy, classification error or usable coverage. For example, the SatVu product specification lists 3.5 m nadir pixel spacing and 15 m CE90 absolute location accuracy: the two values describe different properties. A derived thermal layer also needs separate labeling. NASA specifies Landsat 8 thermal measurements at 100 m, resampled to a 30 m grid; a further 10 m downscaled product is model-derived detail, not native 10 m thermal sensing.

The supplied wildfire material contrasts a 375 m pixel with a 3.5 m pixel. Their nominal nadir areas are 140,625 m^2 and 12.25 m^2, respectively, yielding approximately 11,480 times more samples per unit area for the finer grid. This arithmetic does not imply 11,480 times greater accuracy or a minimum detectable fire equal to one pixel. Detection can exploit a subpixel thermal signal, whereas identification and temperature retrieval have additional constraints. A finer image cannot compensate for the wrong spectral modality, an obscured target or a missed decision window.

Freshness requires the same discipline. A revisit opportunity is not an acquired, cloud-free, licensed and processed observation. Let t^d be decision time and t^a acquisition time. The age of information is $AoI = t^d - t^a$, irrespective of when the map was rendered. A dashboard refresh may retrieve a recent sensor record, an older satellite scene and a static tree inventory simultaneously. Presenting a common screen timestamp would erase the asynchronous nature of the evidence. BeatTheHeat explicitly carries source dates and stale-sensor states, providing an operational precedent for preserving this distinction. Table 3 maps these roles to the three cases; it expresses the synthesis rather than a field-tested ranking of platforms.

Table 3. Complementary sensing roles for the three urban decisions

Case	Satellite contribution	Drone contribution	Hyperlocal contribution
Urban heat	City-wide surface patterns and repeated planning layers	Material-scale thermal inspection and targeted reference campaigns	Air temperature, humidity and local temporal exposure context
Whole car parks	Portfolio inventory, matched-time utilization and land-use context	Facility audit, geometry checks and local count verification	Gate, camera or sensor histories for continuous occupancy
Wildfire / air quality / urban interface	Systematic search, persistence, thermal localization and damage context	Authorized local confirmation and detailed fire/smoke inspection	Local weather, smoke and site observations; incident adjudication

4. Urban heat and climate-responsive design

4.1. What the Cartagena deployment demonstrates

Urban heat products should distinguish surface urban heat islands from canopy-layer air-temperature differences and from human thermal exposure. A hot roof, a shaded footpath and a ventilated open square may have different implications for pedestrians even when viewed in the same orbital scene. A cooling intervention therefore needs a clearly defined outcome. Surface cooling can be evaluated with thermal imagery, while pedestrian comfort also requires local meteorology, shade, wind and radiative conditions. A lower satellite LST is not by itself a measured improvement in individual health or comfort.

The opportunity for cities is a sequence of spatial screening, local design and outcome verification. Broad thermal and land-cover observations can identify neighbourhoods with persistent heat patterns and limited green cover. High-resolution imagery can map trees, roofs and impermeable surfaces. Drone thermography and field instruments can investigate materials and shade around schools, bus stops, care facilities and public squares. The resulting plan can prioritise tree planting, shade structures, cool materials and accessible heat refuges according to both thermal conditions and who uses the space.

AI can help predict temperatures in cloudy periods, downscale broad thermal observations or identify combinations of urban form associated with persistent hot areas. Such outputs are model estimates at the requested grid, not new thermal measurements at that resolution. Validation should hold out entire dates and neighbourhoods, compare with simple baselines and retain uncertainty under unusual heat events. A finely rendered map is useful only if its spatial detail changes an intervention decision and survives independent checks.

Evaluation after intervention should compare treated locations with suitable controls under comparable weather, season and observation geometry. It should report the size and persistence of any change, not simply compare two visually different heat maps. Local sensors supply the temporal context needed to interpret a satellite overpass or drone flight. This creates a service opportunity beyond map production: evidence for design choices and subsequent verification of their effects on public space.

BeatTheHeat provides the applied anchor for the study. Its Cartagena implementation integrates satellite thermal layers, environmental observations, three-dimensional urban morphology, dynamic shade, tree inventories, climate refuges and neighborhood vulnerability. The documented configuration represents more than 16,500 buildings, 36–39 active environmental sensor entities, 120 refuge points and 22 neighborhoods. These are different kinds of coverage, not interchangeable counts of ground instruments. In particular, sensor entities may represent multiple records associated with one installation, and their five-minute application refresh is not a complete specification of physical sampling or network latency.

The thermal service combines Landsat and Sentinel information with machine-learning downscaling and produces a 10 m surface-urban-heat-island (SUHI) planning grid. The case manuscript reports MAE of 0.98 °C and RMSE of 1.25 °C against reference observations for the partner processing component. It does not supply the complete validation sample, reference-instrument description, spatial split or residual dataset required to independently reproduce those errors. They are therefore meaningful reported component results, but not a guarantee for every building or acquisition. The 10 m grid is kept distinct from native thermal resolution.

4.2. Surface temperature is not pedestrian exposure

The most important comparison is physical before it is statistical. Satellite and drone thermography estimate surface-related radiance and temperature; conventional environmental IoT nodes measure local air conditions. A hot roof, a shaded pavement and pedestrian air at the same horizontal coordinate can all have different temperatures without any instrument being wrong. Surface emissivity, viewing geometry and acquisition time also affect comparison. Thermal remote sensing is therefore a spatial screening layer, not a direct replacement for air-temperature or human thermal-comfort measurements (Voogt & Oke, 2003).

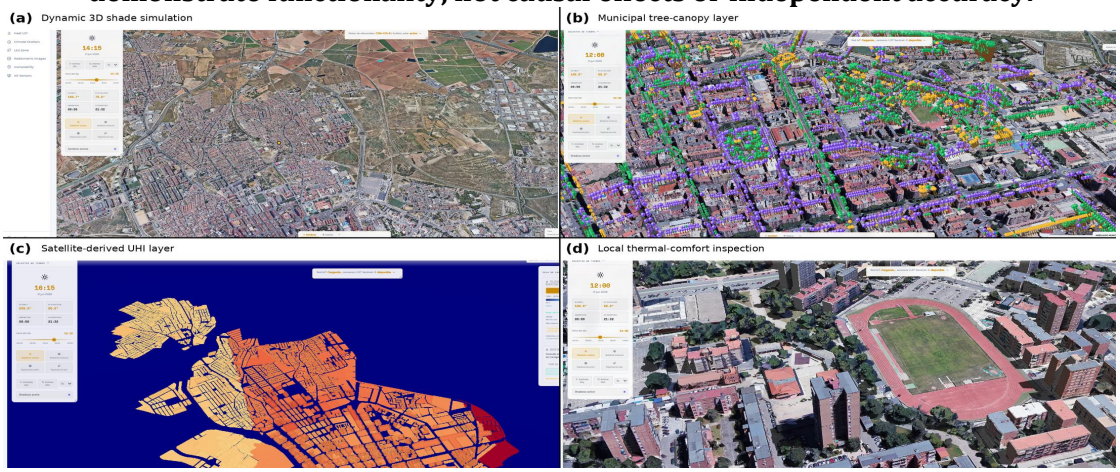
A published paired comparison illustrates the scale trade-off. In an urban-park study using in situ surface reference measurements, Kim et al. (2021) reported RMSE of 3.502 °C for UAV temperature imagery and 7.246 °C for satellite imagery. The result supports the value of fine-scale local imaging in that evaluation, while mixed surfaces, timing and spatial support remain part of the measured discrepancy. It does not establish a universal UAV error or a universal satellite error. Nor can BeatTheHeat's 1.25 °C component result be ranked above the UAV result: the models, references, sites and evaluation protocols differ.

BeatTheHeat integrated field thermography with approximately 2–5 cm sampling at capture distance, but its documentation did not report the synchronized sample size and satellite–thermography correlation needed for a completed independent comparison. It also did not report a full metrological co-location assessment of its IoT network. Table 4 therefore places available errors beside the missing measurements.

Table 4. Thermal evidence: numerical results and comparison boundaries

Evidence	Reported result	Spatial / validation interpretation
BeatTheHeat LST component	MAE 0.98 °C; RMSE 1.25 °C	10 m model-derived grid; partner result; full sample and split unavailable.
Paired park study: UAV	RMSE 3.502 °C	Surface-reference comparison in Kim et al. (2021); local campaign.
Paired park study: satellite	RMSE 7.246 °C	Same study; coarser spatial support is part of the comparison.
BeatTheHeat field thermography	2–5 cm sampling; no matched error reported	Inspection capability; not a completed satellite/UAV validation.
BeatTheHeat environmental IoT	Five-minute application refresh; no co-location error reported	Point air observations; not directly equivalent to surface temperature.

Figure 2. Implemented urban-heat views in the Cartagena digital twin. Panels show (a) dynamic shade, (b) tree canopy, (c) satellite-derived heat and (d) local inspection. Screenshots demonstrate functionality, not causal effects or independent accuracy.



4.3. From a heat map to an urban intervention

The case suggests a transferable planning workflow. First, satellite-derived anomalies identify a block or facility for investigation. Second, the analyst checks whether the hot surface is a roof, pavement, industrial asset or exposed public space. Third, shade geometry and local air observations clarify when and where people may be exposed. Fourth, tree condition, water constraints, neighborhood vulnerability and refuge access inform intervention design. A drone or ground thermographic campaign is commissioned when the uncertainty concerns material-scale conditions that influence the design choice. This workflow is a synthesis of the implemented functions, not a demonstrated reduction in health burden.

The most useful outputs are consequently shade deficits, inspectable surface anomalies, exposure-relevant planning areas and traceable intervention priorities. A percentile-based SUHI score can express how unusual an urban–rural difference is within a city's history, but equal percentiles across cities do not imply equal physical temperatures. Absolute LST and temperature differences must remain available. Similarly, the reduced thermal-comfort estimate in BeatTheHeat remains a screening calculation, not a fully validated UTCI model. These distinctions make replication more credible because a new city can reuse the architecture while recalibrating its thermal baseline and local decision rules.

The strongest direct drone–satellite comparison in the selected evidence concerns Geum River Park in Sejong, South Korea. Kim et al. (2021) studied an approximately 150 m by 300 m area using UAV thermal imagery, Landsat and contact surface thermometers. Seven campaigns were conducted, but cloud or fog excluded two Landsat dates. The reported UAV output grid was 0.09 m and the Landsat output grid 30 m. The latter must be interpreted alongside the instrument's coarser native thermal support.

Table 5. Temperature error on the five common urban-park campaign dates

Campaign date	UAV RMSE (°C)	Landsat RMSE (°C)	Difference (°C)
9 August 2019	2.523	6.502	3.979
28 October 2019	5.298	6.585	1.287
20 March 2020	3.834	6.640	2.806
7 May 2020	2.849	6.385	3.536
18 August 2020	2.788	9.593	6.805
Mean of date RMSEs	3.458	7.141	3.683

Across the five common dates, the calculated arithmetic mean of date-level RMSE is 3.458 °C for the UAV and 7.141 °C for Landsat, a 51.6% reduction for the UAV on this statistic. UAV error is lower on every retained date. The calculation is shown in Table 5; it does not pool individual residuals or estimate statistical significance. Five usable satellite dates out of seven planned campaigns also illustrate why conditional accuracy and usable observation coverage should be reported separately.

This result supports detailed thermal investigation of heterogeneous public space. It does not imply that every drone thermal survey will halve satellite error. Material-dependent emissivity, calibration, atmospheric correction, view angle, georegistration and surface variability can all affect the comparison. Nor does it make ground instruments an error-free third platform: their calibration and representativeness define the quality of the reference. A controlled city evaluation should retain the raw paired observations and report both footprint-matched and material-specific errors.

5. Car parks: satellite portfolios rather than parking-space guidance

5.1. Evidence from satellites and other data sources

Parking offers a useful test of the continuum because the target can be defined at several scales. Satellite measurements can establish where open parking exists and how much is occupied at acquisition time. Drones can inspect individual facilities. Fixed cameras and per-space sensors can track local changes. Table 6 retains the original metrics of representative studies so that their evidential scope remains visible.

Table 6. Published parking and vehicle results with their measurement units

Study and setting	Task and reference	Reported result	Comparison limit
Stopic et al. (2023), Alkmaar	0.5 m Pléiades; occupancy of mapped spaces; synchronous field reference	F1 0.82; 0.85 after canopy exclusion	Higher conditional score excludes some spaces.
Sheehan et al. (2023), Barcelona	WorldView and YOLOv3; vehicle detections against annotations	Precision 0.69; recall 0.79	Vehicle detection, not space occupancy or continuous flow.
Peraza-Garzón et al. (2026), university car park	UAV and YOLOv11; 75-image test set after dataset augmentation	Precision 0.94; recall 0.92; F1 0.93	Source-image and flight independence need checking.
Encío et al. (2023), PKLot	Fixed-camera space classification; stated training and test split	Mean accuracy 99.14%	Different task and denominator from detection F1.
Drouyer & de Franchis (2020)	Sentinel-1	MAE 2.56 percentage points; r = 0.96	21-scene, one-site calibration; application to 250 facilities, not 250-site validation.
Pléiades; Stopic et al. (2023)	Space classification from 0.5 m imagery	F1 ≈ 0.82; ≈ 0.85 without canopy	Aggregate facility error is a different metric.
Lou et al. (2019)	IoT parking spots as the Libelium parking sensors	Detection accuracy up to 99.62%; 1 Hz sampling	Instrumented local parking events; deployment and aggregation define facility coverage.

The Alkmaar evaluation is particularly relevant to land allocation because its reference observations were collected at the satellite acquisition time. Its F1 increased from 0.82 to 0.85

when tree-obscured spaces were excluded using vegetation information (Stopic et al., 2023). The improvement supports explicit observability masks. It does not establish that the hidden spaces were empty or that the whole city was observed with the higher score. Barcelona demonstrates a broader vehicle-detection application, with precision 0.69 and recall 0.79, and documents the consequences of narrow streets, shadow, vegetation and buildings (Sheehan et al., 2023).

The UAV study reports 0.94 precision, 0.92 recall and 0.93 F1, with 35–38 frames per second on a Tesla T4. However, the 750-image dataset was expanded from 30 selected source images, and 75 images formed the test set. Independence at the original flight or scene level therefore needs scrutiny before interpreting the score as deployment generalisation. We use its results table value of 0.96 mAP@0.5; an introductory passage instead states 0.97 (Peraza-Garzón et al., 2026). Figure 2 reproduces an actual published output, including overlapping detections, to make the remaining aggregation problem visible.

Fixed-camera evidence gives a different comparison. Encío et al. (2023) report 99.14% mean occupancy classification accuracy on PKLot under their stated dataset split. This is classification of predefined spaces using camera imagery, not satellite vehicle detection or lot-level percentage error. It demonstrates strong local feasibility without establishing equivalent performance on unobserved streets or new cameras. Likewise, continuous updates require a deployed communications and maintenance service; they do not follow automatically from image-classification accuracy.

A separate scalable route uses SAR as an aggregate indicator. Drouyer and de Franchis (2020) evaluated a parking backscatter relationship against a camera at one site and extended their analysis to 250 large French retail parking areas. This is evidence for broad activity monitoring, not individual-space recognition. Satellite, drone and ground results thus support different products. They cannot establish a single accuracy-versus-area curve because the tasks, datasets and observation intervals differ. Their Sentinel-1 backscatter method was calibrated against occupancy estimates from 21 scenes at one camera-observed site. The best configuration reported a Pearson correlation of +0.96 and MAE of 2.56 percentage points. Parameters were optimized on that small sample, so the error should not be presented as independent held-out performance.

For a physical IoT comparison, Lou et al. (2019) evaluated magnetic sensing supplemented by wireless signal-strength information in a real roadside parking setting. They reported vehicle-detection accuracy up to 99.62% with 1 Hz magnetic sampling. The result demonstrates local event sensing under the study's conditions, not the performance of a particular commercial product or an entire city. Its relevance here is the temporal reference it can provide when aggregated to facility level. The satellite service observes uninstrumented territory, whereas continuous local sensing requires deployment at the locations it represents.

5.2. Planning opportunities and a transferable service design

The most immediate opportunity is a repeated inventory of open surface parking linked to land use. Cities can identify underused paved areas, test park-and-ride assumptions, investigate event overspill and prioritise sites for trees, permeable surfaces or alternative uses. These decisions usually tolerate a longer observation horizon than driver guidance. Their evidence should nevertheless cover representative weekdays, weekends, seasons and events. A single acquisition is inadequate to justify permanent removal of capacity or to estimate turnover.

A practical service would first establish parking polygons and capacity from imagery and municipal records. Selected acquisitions would estimate occupancy and observable capacity. A local instrumented sample would then quantify errors and temporal representativeness, while drones would resolve ambiguous layouts or canopy effects where permissions allow. The final product would identify candidate sites for investigation and explain uncertainty. Decisions would combine these observations with accessibility, loading needs, disabled parking, resident feedback and public-transport access.

In a Mediterranean port city such as Cartagena, this design could connect peripheral car parks, waterfront event areas and access corridors with heat and flood exposure. This is a transfer opportunity proposed here, not a claim of newly measured local parking performance. The same paved area might be lightly used outside events, excessively hot during summer and important for emergency assembly. The advantage of a common observation portfolio is that the city can assess these competing functions together instead of procuring unrelated maps for each department.

The car-park case concerns open facilities, park-and-ride networks, event areas and logistics staging grounds. Its principal questions are spatial: where capacity exists, which facilities appear saturated at comparable times, where informal overflow occurs and which paved areas warrant land-use review. A satellite observation is an intermittent facility snapshot, not a reliable instruction that an individual bay remains vacant when a driver arrives. Covered and underground interiors are outside overhead optical visibility; trees, buildings and shadows can make even an open facility partly unobservable. The evidence is strongest when the output and denominator are explicit. For a valid observed subarea, occupancy is the estimated number of parked vehicles divided by its usable capacity. Capacity may come from an independently maintained facility inventory rather than a predicted collection of individual spaces. Where capacity is unknown or unstable, the defensible output is a count, density or activity category, not an exact occupancy percentage. Unobservable area is reported separately and is never interpreted as empty.

The proposed facility product combines a versioned car-park polygon, capacity record, visible vehicle estimates or calibrated proxy, observation time, observed-capacity fraction and uncertainty. The inference pipeline should preserve original imagery, use suitable tiling for small objects, remove duplicate detections and separate parked vehicles from adjacent roads. A count correction learned at one site must not be transferred automatically to another morphology or sensor. For a future evaluation, facility-level occupancy error in percentage points and count bias are more relevant outcomes than detector mAP alone.

The planning opportunity extends beyond parking management. A consistently underused paved facility may be a candidate for shared use, shade, permeable surfaces or redevelopment; a saturated park-and-ride may warrant investigation of access and transit integration. These are hypotheses requiring local corroboration, not causal conclusions from one image. Combining utilization with heat and shade layers creates a common evidence base for transport and climate departments. Local sensors retain responsibility for live availability, while orbital observations make portfolio-wide review possible without first installing a device in every bay.

6. Wildfire and the urban interface: detection, confirmation and consequence

6.1. Systematic observation provides the surveillance backbone

Urban wildfire intelligence includes the wildland–urban interface, access corridors, industrial boundaries, exposed infrastructure and smoke reaching residential areas. Its tasks must remain separate: detecting a thermal candidate, confirming an actionable event, locating the active source, characterizing evolution and assessing consequences. An apparent gain in one task does not validate the others. The supplied wildfire framework therefore treats thermal, visible optical and synthetic-aperture radar (SAR) products as non-equivalent evidence sources rather than interchangeable high-resolution tests.

Recent European evidence is quantitatively substantial. Xu et al. (2026) evaluated Meteosat Third Generation's Flexible Combined Imager (FCI). Relative to MODIS and VIIRS reference comparisons, FCI omission rates were 38% and 68%, compared with 83% and 89% for SEVIRI; corresponding FCI commission rates were 12% and 10%. Matched MODIS fire-radiative-power retrievals showed $R^2 = 0.97$. Selected Portuguese and Greek events were detected up to four hours earlier than SEVIRI or VIIRS and two hours earlier than MODIS. These are study-specific comparisons, not guaranteed warning lead times for all fires. Their importance is that systematic

temporal sampling and improved fire sensitivity can create value without resolving every source at metre scale.

A practical workflow maintains systematic open observations for regional screening and persistence, then requests a more detailed thermal or optical observation when localization or context could change the decision. SAR can contribute under cloud or darkness to surface-change, access and damage questions, but it does not directly measure fire temperature or radiative power. Local weather and incident records constrain interpretation. Where an immediate threat is credible, imagery acquisition must not delay authorized emergency action; decision deadlines are a hard requirement, not a financial variable that can override safety.

Wildfire is an urban-planning concern because settlements, transport corridors and critical infrastructure interact with surrounding vegetation. Observation supports fuel and exposure mapping, incident assessment, smoke management and recovery planning. These tasks have different deadlines and should not be collapsed into a single early-warning claim. A useful historical burn map may arrive too late to inform initial response, whereas a rapid thermal candidate may be sufficient to trigger local verification without a detailed perimeter.

Pereira et al. (2021) provide real AI validation using Landsat 8 imagery. Their best combined approach achieved 87.2% precision and 92.4% recall against manually annotated active-fire masks; the harmonic mean of these rounded values is approximately 89.7%. The manual dataset includes a California scene acquired on 8 September 2020, providing an identifiable real-event example. These are pixel-segmentation results, not event-level detection probability or emergency-service false-alarm rates. The distinction matters because one incident can produce many positive pixels, while a small or obscured incident may generate none.

The acquisition portfolio must follow the physical question. VIIRS active-fire observations support broad thermal-anomaly surveillance (NASA Earthdata). Detailed thermal imagery can investigate heat patterns; optical imagery can show smoke, access routes and exposed buildings. SAR provides complementary information on structure and surface change through cloud, but does not directly measure flame temperature or fire radiative power. A SAR acquisition should therefore not be scored as if it were an interchangeable thermal confirmation test.

The commercial examples supplied with the starting material are useful as an acquisition portfolio. Current documentation lists 3.5 m thermal sampling for SatVu, 30 m thermal products for constellr, 35 cm Gen-3 optical products for BlackSky, and several distinct ICEYE SAR modes (UP42). These are modality-specific specifications. They do not verify a common sensitivity, guaranteed acquisition at a particular site or delivery before a fire-response deadline.

Reducing nominal square-pixel spacing from 375 m to 3.5 m gives an area ratio of approximately 11,480. That geometric contrast can reduce spatial ambiguity, but it does not imply 11,480 times better accuracy or a proportional reduction in false alarms. Hot subpixel targets can be detectable within a much larger thermal pixel, and spectral response, saturation and atmospheric effects also matter. Municipal services should ask whether additional evidence changes a decision before its deadline, rather than equating finer pixels with operational certainty.

A value-of-information model provides an explicit acquisition rule. Let p be the current probability of an actionable incident, C_R the assumed response cost and C_FN the consequence of not responding when the incident is real. In a simplified two-action model, responding has loss C_R and not responding has expected loss pC_FN. The baseline loss is therefore min(C_R, pC_FN), and the action threshold is C_R/C_FN. These are analytical assumptions, not an emergency dispatch protocol.

A candidate observation changes the posterior probability. Its expected value of sample information (EVSI) is the baseline loss minus the probability-weighted minimum loss after the possible observation outcomes. Let q be usable-delivery probability, C_full fully loaded acquisition and processing cost, τ delay and λ an assumed loss per fire-hour. A candidate is economically admissible in this simplified model only if $C_full \leq qEVSI - p\lambda\tau$ and the modality, coverage, quality and deadline requirements are satisfied. Sensitivity and specificity must be estimated for the actual target decision; assigning the same assumed values to different products is only a scenario probe.

Table 7 uses explicit scene-area scenarios to expose the scaling issue independently of changing quotations. At 500 km^2, even perfect packing requires at least 32 scenes of 16 km^2 or 20 scenes of 25 km^2. At 225 km^2 per scene the bound is three scenes. These calculations concern area coverage only. They do not imply equivalent thermal information, geometry, timing or scene availability. Any operational mosaic requires a further assessment of overlap, missing coverage and the time span between its earliest and latest acquisitions.

Table 7. Area-only lower bounds for hypothetical scene portfolios

Requested area (km^2)	16 km^2 scenes	25 km^2 scenes	225 km^2 scenes
25	2	1	1
100	7	4	1
500	32	20	3

The framework favours targeted acquisition where an additional observation can alter an uncertain decision. At low incident probability, continued monitoring may already be appropriate; at high probability, waiting may add little information and consume valuable time. For spatial planning, the same logic applies over longer horizons: detailed imagery may be justified to resolve uncertain vegetation setbacks or infrastructure exposure, while broad public data may suffice for initial regional screening. Actual procurement should compare cost per accepted, timely decision unit rather than price per scene.

Air quality connects the fire, mobility and urban-form applications. Satellite aerosol or atmospheric-column products describe regional conditions, while ground stations and IoT sensors characterise local concentrations. They are physically different observables. A surface PM2.5 estimate derived from orbital data requires a model linking atmospheric observations to surface conditions, usually with meteorology, land use and ground measurements. Fine optical imagery of cars does not directly measure the concentration that a pedestrian breathes.

Wei et al. (2023) demonstrate the scale achievable through this fusion. Their global daily 1 km PM2.5 model reports sample-based cross-validation R^2 of 0.91 and RMSE of 9.2 $\mu g/m^3$. Station-based validation reports R^2 of 0.87 and RMSE of 10.94 $\mu g/m^3$, making the cost of spatial transfer visible. These figures refer to a modelled surface product, not direct 1 km PM2.5 sensing. They should not be compared as if they represented the same experiment as local sensor calibration.

Ground sensing also requires calibration. Barkjohn et al. (2021) reduced PurpleAir PM2.5 RMSE from approximately 8 to 3 $\mu g/m^3$ using a US-wide correction, including relative humidity, against reference or equivalent monitors. This result demonstrates the importance of calibration; it is not a universal specification for every low-cost sensor or pollution mixture. A city portfolio should evaluate satellite-assisted maps and local devices against compatible withheld reference observations, with separate results for ordinary conditions and smoke, dust or other unusual episodes.

Planning opportunities include locating monitoring stations, identifying neighbourhoods needing closer investigation, assessing regional smoke transport and combining traffic activity with exposure modelling. School streets, low-emission measures and green infrastructure can be prioritised using several layers, but their effects require evaluation beyond a coincident change in imagery. Appropriate weather controls, source information and ground measurements are needed to distinguish intervention effects from regional transport or seasonal change. This creates demand for services that join spatial reach with accountable local evidence.

6.2. What drone and IoT evaluations add

Pesonen et al. (2024) reported 63.3% mean intersection over union for UAV wildfire segmentation on their annotated evaluation dataset and approximately 25 frames per second on a UAV-carried Jetson Orin NX. They also demonstrated the system during forest-burning events.

Offline segmentation quality, embedded processing speed and field deployment are distinct results: together they support practical local inspection, but they do not establish regional ignition-detection sensitivity or a full operational dispatch time. The European research example illustrates the drone's role as a targeted, mobile evidence source rather than a substitute for persistent wide-area search.

Local sensing is likewise conditional. Guth et al. (2026) evaluated low-cost sensor measurements of prescribed-fire smoke, reporting 61% detection accuracy and a 3% false-positive rate for their random-forest approach, compared with 51% and 31% for the alternative algorithm reported in the study. These results concern smoke-plume classification in the evaluated setting, not detection of all uncontrolled wildfire ignitions. A local network's spatial reach depends on transport from the source to its sensors; silence at a node is not evidence that the wider territory is fire-free. Table 8 summarizes the non-equivalent but complementary results.

Table 8. Wildfire evidence and the limits of cross-modal comparison

Observation	Reported evidence	Decision meaning
Meteosat FCI; Xu et al. (2026)	Omission 38% / 68% and commission 12% / 10% versus MODIS / VIIRS	Wide-area active-fire comparison; reference and matching rules matter.
UAV; Pesonen et al. (2024)	63.3% mIoU; about 25 frames/s	Segmentation benchmark and onboard processing; field demonstrations reported.
Local sensors; Guth et al. (2026)	61% detection accuracy; 3% false-positive rate	Prescribed-fire smoke classification; not regional ignition sensitivity.
Aistech and other commercial thermal products	Specifications and mission development; no common benchmark here	Potential targeted localization; provider-specific fire accuracy is not established by this synthesis.

The wildfire case makes the economic boundary especially clear. The value-of-information framework in the supplied study compares the expected reduction in decision loss with fully loaded acquisition cost and the consequence of waiting. Its illustrative sensitivity and specificity of 0.95 and usable-delivery probability of 0.80 are model assumptions, not measured performance of Aistech, SatVu, constellr, BlackSky or ICEYE. The resulting monetary thresholds are therefore retained as methodological context rather than reproduced as observed savings or operational response advice.

A useful implementation records the event, the source asset and the decision separately. It preserves the original uncertainty footprint, the question asked of the new observation, the deadline, the delivered usable area and the final field adjudication. Such a record can establish whether a detailed image changed an action, whether a drone visit was still required and whether a recurring industrial source was correctly classified. The future training asset is therefore not just a collection of images, but a governed set of event–evidence–outcome relationships.

7. European thermal missions and the VLEO opportunity

7.1. An expanding portfolio, not one universal sensor

The European opportunity includes public systematic missions and commercial capabilities with different sensing physics. Table 9 records representative enabling systems using official information checked at the review date. The purpose is to make procurement questions explicit rather than to rank providers. European origin, present orbital status, operational data availability and future fleet objectives are separate attributes. Likewise, a provider's

constellation-average revisit is not a guarantee that the required city will receive a usable product before its deadline.

Table 9. Representative European capabilities and development boundaries

System / origin	Documented capability or goal	Urban opportunity / boundary
Meteosat FCI / Europe	10-minute full-disc observation; dedicated active-fire evaluation	Regional persistence and early fire evidence; not building-scale thermal mapping.
Aistech Hydra / Spain	Provider lists Hydra 1–3 in orbit; 7 km swath and 35 km^2 scenes	Targeted thermal services; orbital status is not independent fire-performance validation.
SatVu / United Kingdom	3.5 m nadir spacing; nominal 16 km^2 scene; day/night MWIR	Local thermal anomalies and source context; location accuracy, cloud and timing remain separate.
constellr HiVE / Germany	Two satellites delivering data; 30 m native thermal, 10 m sharpened; 1.5-day mean revisit	Urban surface patterns; sharpening and average revisit are not native detail or an SLA.
ICEYE / Finland	Multiple SAR imaging modes trade detail against coverage	Access, surface change and damage context; not direct thermal confirmation.
Aistech-led VLEO initiative / Spain	Development target near 200 km with thermal imaging, propulsion and AI	Emerging opportunity; not a demonstrated operational VLEO fire-warning service.

7.2. Why lower orbit can matter for urban planning

The geometric opportunity is straightforward: for an otherwise fixed nadir optical configuration, reducing altitude reduces the ground distance represented by an angular sample. This can improve separation of small urban features or permit different instrument-size trade-offs. The corresponding service benefit is not automatic. A narrower ground footprint, viewing angle, attitude stability, radiometric calibration, drag management and lifetime can affect both usable detail and economics. VLEO is therefore evaluated here as a development pathway that may improve selected observations, not as a synonym for continuous drone-like city monitoring.

Promising future applications include isolating thermal behavior within complex facilities, improving the interpretation of mixed urban surfaces, and inspecting spatially compact anomalies selected by regional systems. Their maturity should be assessed through delivered scenes, calibrated measurements, real acquisition opportunities and end-to-end latency. The city should procure the decision product it needs, not an orbital altitude. A future VLEO thermal product might complement an existing heat service or refine a wildfire candidate; it would not eliminate the need for local air measurements, ground confirmation or independent validation.

VLEO can create opportunities where additional ground detail changes the planning question: distinguishing narrow planted strips, improving small-asset inventories or identifying geometry that coarser imagery merges. Dedicated thermal constellations can add temporal and spatial observations useful for investigating surface patterns. Drone autonomy and edge AI may reduce the effort needed for repeated site inspection. These developments should be assessed separately because improvement in one capability does not establish operational maturity in the others.

The most promising innovation is adaptive observation across the portfolio. A broad anomaly map can select a limited area for further inspection, and the new evidence can update both the urban indicator and future acquisition priorities. Such a service requires calibrated uncertainty; otherwise it may repeatedly inspect already well-observed districts while ignoring areas where the model fails silently. Geographic diversity matters. The VME benchmark explicitly broadens

satellite vehicle examples across Middle Eastern and other cities, illustrating why familiar benchmark performance cannot replace geographically representative validation (Al-Emadi et al., 2025).

Public value also depends on how observation affects the use of urban space. A cooling investment should account for access and vulnerability as well as the hottest pixels. A parking intervention should consider users whose needs are poorly represented by occupancy alone. Urban observation should therefore be linked to public consultation and administrative evidence, with aggregate outputs appropriate to the decision. The observation portfolio can reveal unequal conditions; it cannot independently decide how competing needs should be prioritised.

7.3. Services organised around planning outcomes

The opportunity extends beyond the four principal applications. A city can reuse georeferenced observations across departments while retaining task-specific validation. Parking and heat layers can inform the conversion of paved space; vegetation and thermal information can inform water-sensitive landscape management; SAR and optical observations can support flood and infrastructure investigations. Table 10 maps additional opportunities to concrete decisions and complementary local evidence. These are proposed transfer pathways, not additional validated deployments in this study.

Table 10. Additional urban opportunities and the evidence needed locally

Planning opportunity	Remote sensing contribution	Local evidence and decision
Flood resilience	SAR flood extent, terrain and optical impervious-cover mapping	Water levels and site surveys for drainage, access and retention planning.
Urban green infrastructure	Canopy extent, vegetation condition and thermal context	Tree inspections, soil moisture and irrigation records for maintenance and planting.
Roofs and energy	Roof geometry and thermal anomalies	Building and energy audits before selecting retrofit or solar interventions.
Infrastructure condition	SAR deformation time series and optical change	Survey or engineering inspection to investigate movement and prioritise repairs.
Coastal and port planning	Shoreline, land use and surface-change observations	Tide, bathymetric and operational evidence for exposure and access decisions.
Urban growth and land use	Repeated built-form and construction mapping	Planning records and field checks for infrastructure demand and land allocation.
Waste and brownfields	Visible surface change and candidate dumping areas	Site inspection and material sampling before enforcement or remediation.
Heat-resilient public space	Combined parking, vegetation and surface-temperature layers	Use patterns and comfort measurements for shade, permeable surfaces and refuges.

For EO providers, the useful unit of service is the decision product. A planner may need a defensible list of candidate cooling interventions, a validated parking inventory or an updated exposure map. The value lies in converting observations into consistent geographic units with visible uncertainty, not in delivering an increasingly large collection of images. A common spatial

foundation can reduce duplicate acquisition and processing, but each application still needs its own observable, quality checks and validation records.

A plausible service model combines periodic baseline monitoring with targeted investigations and verification after intervention. The baseline can use open archives and selected commercial acquisitions. Investigations can add higher-detail imagery, drone surveys or temporary instruments. Verification compares subsequent observations with the earlier baseline using a declared method. This arrangement allows a municipality to expand coverage without initially installing permanent infrastructure everywhere, while retaining local instruments where temporal continuity or reference quality is essential.

Commercial viability should be assessed through actual municipal workloads. Relevant quantities include staff time saved in preparing an inventory, the proportion of sites that can be assessed remotely, the number requiring field review, the rate of successful acquisitions and the cost of maintaining the service. Savings, avoided losses and environmental benefits should be measured in pilots rather than assumed from the imagery specification. A successful technical demonstration can still fail as a service if manual review, tasking delays or licensing prevent routine use.

Air quality provides a natural extension because heat, transport activity, industrial land use and wildfire smoke can affect the same places while requiring different measurements. Sentinel-5P/TROPOMI nitrogen-dioxide products include tropospheric-column information and retrieval-quality variables. A column quantity must not be labeled as street-level concentration without a validated conversion model and suitable atmospheric information (TROPOMI). The opportunity is to combine regional patterns with calibrated local observations and explicit modeling, not to draw a high-resolution-looking regulatory map from an unrelated image grid.

BeatTheHeat already incorporates pollutant observations alongside thermal screening and retains flagged or stale sensor states. Its joint heat–air-quality alerts demonstrate integration functionality, not an epidemiologically validated combined-risk score. A transferable service could use orbital information to select underobserved areas, local sensors to characterize temporal exposure, and targeted field campaigns to evaluate model bias. The same provenance requirements apply to both pollutant and temperature products: observed, modeled and imputed values should remain distinguishable.

Table 11 sets out an opportunity agenda derived from the three cases. These are proposed service directions, not additional validated deployments. Their common structure is a repeatable spatial question, a limited set of observable variables and a local verification path. This is a more actionable basis for market development than an unsupported assertion that all urban functions will benefit equally from finer imagery. It also identifies potential users and success measures without inventing a market-size forecast.

Table 11. Opportunity agenda for city-oriented remote-sensing services

Service direction	Potential urban user and decision	Evidence needed before outcome claims
Heat and shade planning	Climate and public-space teams: prioritize exposed facilities and routes	Matched thermal references, shade checks and intervention baselines.
Car-park land-use review	Transport and planning teams: utilization, shared use and redevelopment screening	Facility-level counts, capacity records and matched observation windows.
Urban-interface resilience	Civil protection and infrastructure operators: verify threats and review access	Incident adjudication, latency logs and decision-specific validation.
Air-quality spatial screening	Environmental teams: identify monitoring gaps and investigate co-exposure	Reference co-location and validated column-to-surface modeling where used.
Greening and water-sensitive design	Urban forestry and drainage teams: inspect canopy and impervious-area priorities	Local inventory, tree condition, water and drainage evidence.
Asset and energy screening	Municipal property teams: target roof or facility inspections	Calibrated thermal surveys and engineering assessment; no direct energy-loss inference.

A particularly useful cross-case object is the open car park. Its utilization can inform transport planning, its surface and shade can inform heat adaptation, and its access and

surroundings can matter for emergency staging. A common facility geometry enables departments to compare these dimensions without assuming that an underused site should automatically be removed or that a hot site requires one particular intervention. The resulting service can support a structured option appraisal: retain, share, shade, redesign or investigate. Decision ownership and local evidence remain essential.

A second opportunity is verification as a service. Municipalities may not need permanent specialist analysis for every orbital acquisition, but they do need an auditable route from a flagged location to a dependable interpretation. A provider-neutral service could coordinate a targeted drone mission, a calibrated ground measurement or expert review, then write the result back to the city's evidence record. The value proposition is reduction of unresolved uncertainty within a deadline. It must be tested through documented decisions, not assumed from the visual quality of an interface.

8. Scalability, economics and a reproducible validation pathway

8.1. Normalize the service before comparing the cost

Scalability has at least three dimensions: geographic extent, temporal continuity and the number of decision units served. Satellites can add spatial coverage without a corresponding installation at every location; drones can concentrate detail but require mission time; fixed IoT provides continuity where infrastructure exists. These structural differences do not establish which option has the lowest equivalent-service cost. A monthly portfolio audit and a continuous occupancy service have different outputs. The original mobility and wildfire studies both warn against comparing one scene price with an operated multiyear sensing network.

Coverage arithmetic exposes one limitation. With a hypothetical compact 100 km^2 area and ideal packing, a 25 km^2 scene requires at least four acquisitions, a 16 km^2 scene at least seven, and a 35 km^2 scene at least three. These are geometric lower bounds calculated here from representative nominal footprints, not feasible tasking plans. Real arrangements must account for scene shape, overlap, clouds, off-nadir viewing and disconnected facilities. Conversely, one scene can serve several planning questions when the licence permits reuse. Coverage economics therefore depend on the useful area and decisions served, not only nominal image price (UP42; SatVu; Aistech Space).

A service-normalized cost measure is total acquisition, deployment, operation, processing and review expenditure divided by the number of decision units whose coverage, error and age requirements are actually met. Failed acquisitions and periods of sensor outage remain in the cost numerator. Useful comparison units include a quality-qualified facility audit, a neighborhood thermal assessment or an adjudicated fire candidate. Public scene tariffs are deliberately not treated as stable service quotations in this paper. Local feasibility, licences, review and integration can dominate the final cost.

8.2. Acquire additional evidence only when it can improve the decision

The supplied wildfire study formalizes a useful selection rule. For product j, uncertainty area U and actionable-event probability p, additional acquisition is admissible only when mandatory quality, rights, modality and deadline conditions are satisfied and its fully loaded cost does not exceed expected usable information value minus delay loss (Cuevas Martínez et al.):

$$C_j^{full}(U) \leq q_j(U)\, EVSI_j(p) - p\,\lambda\,\tau_j(U) \quad (1)$$

Here q is the probability of usable delivery, EVSI is expected reduction in decision loss, λ is incremental loss per hour conditional on the event, and τ is end-to-end decision latency. These quantities require local estimation; an uncalibrated model confidence is not a valid event probability. For non-emergency planning, the same reasoning can compare whether additional detail is likely to change an investment choice. The equation is adopted from the supplied framework, not presented as a newly validated economic result. It complements rather than replaces institutional safety procedures.

8.3. A common prospective benchmark

A defensible next experiment would pre-register decision units, acceptable errors and required freshness before selecting products. The heat campaign would compare synchronized radiometric surface references, UAV observations and satellite products after aggregating to common spatial support; air-temperature validation would remain a separate co-location task. The car-park campaign would lock facility geometry and capacity, obtain synchronized manual or continuous reference counts, and evaluate all methods using facility occupancy error and observed fraction. The wildfire campaign would retain both actionable and non-actionable candidates and link observations to independent incident adjudication, avoiding a dataset composed only of successfully detected fires.

All three campaigns should separate training, calibration and testing geographically and temporally; report missing observations and abstention; and resample uncertainty at site, scene or event level rather than treating adjacent pixels as independent trials. Accuracy must be accompanied by usable-coverage yield, decision-time age, human review, end-to-end cost and the action before and after the evidence. This protocol is prospective. Its purpose is to convert today's heterogeneous but informative results into tomorrow's service-level comparison without pretending that the required common dataset already exists.

8.4. A common validation design for deployment

A future field evaluation should compare complete service configurations on the same urban decisions. It should include open and shaded parking, heterogeneous public-space materials, urban-fringe exposures and distinct pollution settings. Satellite-only, drone-only and ground-based configurations should be evaluated alongside their fusion. Acquisition should be synchronised where the target changes quickly, and withheld sites and dates should be defined before model tuning. Data augmentation must remain within the training partition of its original acquisition.

For parking, the evaluation should report precision and recall for detections, occupancy-fraction error for lots, unknown-space rate and usable capacity. For heat, it should report surface-temperature bias and RMSE, reference-instrument uncertainty, material strata and footprint-matched comparison. For wildfire, pixel metrics should be accompanied by independently adjudicated incident detection, false alerts per area-time and detection-to-decision delay. For air quality, station-held-out errors and episode-specific performance should accompany ordinary sample-level validation. These quantities answer different questions and should retain their units.

Scalability should then be measured under the same area and time requirements. The evaluation should count requested acquisitions, completed acquisitions, rejected observations, usable area, end-to-end delivery and human review. Fully loaded costs should include acquisition, processing, storage, communications, calibration, maintenance and local verification. A comparison that omits crew time for drones or repeated installation for local sensors can misstate the trade-off just as much as one that ignores cloud and tasking failures for satellites.

Acceptance thresholds should be set with the municipal decision owner. A strategic screening map can tolerate uncertainty that would be unacceptable for direct parking guidance or a consequential response decision. The thresholds should specify when a system can provide an indicator, when it must mark the result uncertain and when it must request another source. This allows the observation portfolio to expand progressively without converting every experimental feature into an operational claim.

8.5. Provenance, interoperability and limitations

Every accepted observation should retain its origin, acquisition time, geometry, processing level, quality masks and permitted uses. A derived indicator should additionally retain the model version, reference-data version, uncertainty and aggregation rule. The SpatioTemporal Asset Catalog specification provides a common structure for describing geospatial assets and their spatial and temporal context (STAC Community). Municipal entities and decision records can

link to those assets without flattening away the differences between a satellite acquisition, a drone survey and a point measurement.

This record supports both reproducibility and service management. When two indicators disagree, the city should be able to determine whether they concern different times, surfaces, spatial footprints or models. When a provider changes processing, the affected series should be traceable. Image rights, permitted downstream reuse and retention also need to travel with the product. The cost of these controls should be included in operational evaluation because a map that cannot be audited or reused has a different value from one that can.

The principal limitation is the absence of a single co-located field dataset spanning satellites, drones and IoT across all four applications. The temperature analysis is a secondary calculation from published aggregate errors; the parking studies use heterogeneous tasks and splits; fire segmentation is not an incident-response evaluation; and pollution-model validation is not a comparison of interchangeable sensors. The scene-area calculations are ideal lower bounds. Provider documentation establishes described capabilities rather than independently verified urban service levels.

9. Discussion and conclusions

Remote sensing can expand the spatial evidence available to smart cities while drones and hyperlocal instruments provide detail, continuity and reference measurements. Real parking, thermal, fire and pollution evaluations demonstrate that useful capabilities already exist. The five-date thermal reanalysis shows a substantial local drone advantage over Landsat on a clearly defined error statistic, while the heterogeneous parking results show why broad modality rankings remain unjustified. Accurate local observation and scalable territorial coverage are complementary design requirements.

The evidence supports a complementary allocation of roles, not a general claim that either satellites or IoT are more accurate. BeatTheHeat shows how an implemented city environment can preserve satellite, point, geometry and social evidence in one planning workflow. The paired thermal study establishes a local UAV advantage under its conditions. Car-park research demonstrates both wide-area aggregate processing and high-frequency local detection, but on different validation populations. Wildfire research shows that systematic European observation can improve detection while UAV and local sensor studies address more localized tasks. None supports a single pooled accuracy percentage.

Three limitations constrain the conclusions. First, this is an applied synthesis of heterogeneous evidence, not a new simultaneous three-platform experiment. Second, BeatTheHeat's reported component errors cannot be independently reconstructed from the supplied manuscript, and the available case does not establish causal cooling, health or financial benefits. Third, commercial and VLEO capabilities evolve; launch, commissioning, available data, calibrated performance and contracted service are different maturity milestones. The integrated cross-case architecture is proposed, whereas the documented urban-heat deployment supplies the relevant-environment implementation evidence. A formal readiness assessment of every proposed extension has not been performed.

The opportunity is to organise acquisition around an urban decision: map broadly, investigate selected uncertainty and verify the resulting intervention. VLEO may extend the detail available from orbit, but its value must be assessed through delivered evidence, usable coverage and reliable timing. For cities and the remote sensing market, the strongest pathway is a portfolio of validated decision products that connects parking, cooling, environmental exposure and infrastructure planning while preserving the meaning and limits of each measurement. Remote sensing can extend the spatial reach of city planning to facilities and neighborhoods where permanent instrumentation is absent. Drones can resolve selected local questions and provide targeted verification. Hyperlocal IoT can preserve temporal continuity and the environmental variables that orbital imagery does not directly measure. European thermal missions and VLEO development can enrich this portfolio when delivered performance matches the intended decision. Market development should therefore emphasize reliable urban

services, qualified facility audits, heat-intervention evidence and adjudicated threat information, rather than a race to the smallest pixel.

From hyperlocal with IoT to the sky with drones and satellites brings a new paradigm of sensing and new opportunities for AI-driven urban planning. A trustworthy urban intelligence service states what was measured, where and when it was observed, how much of the decision area remained unseen, and what further verification is needed. Under those conditions, satellite imagery is not merely another layer in a digital twin: it becomes a scalable means of discovering, comparing and acting on urban opportunities while keeping local evidence and public accountability in the loop.

References

Aistech Space. (2025, October 2). *Satellite-based detection of forest fires*. https://aistechspace.com/satellite-based-detection-of-forest-fires/

Aistech Space. (2026, January 12). *Aistech Space launches the Hydra-2 satellite*. https://aistechspace.com/aistech-space-launches-the-hydra-2-satellite/

Al-Emadi, N., Weber, I., Yang, Y., & Ofli, F. (2025). VME: A satellite imagery dataset and benchmark for detecting vehicles in the Middle East and beyond. *Scientific Data, 12*, 500. https://doi.org/10.1038/s41597-025-04567-y

Barkjohn, K. K., Gantt, B., & Clements, A. L. (2021). Development and application of a United States-wide correction for PM2.5 data collected with the PurpleAir sensor. *Atmospheric Measurement Techniques, 14*, 4617–4637. https://doi.org/10.5194/amt-14-4617-2021

Cuevas Martínez, I., Jara, A. J., & Fernández-Breis, J. T. *BeatTheHeat: A data-space-enabled urban digital twin for multi-scale heat exposure, microclimate design, and equity-aware governance*. ElSevier Internet of Things Journal, 2026.

Cuevas Martínez, I., Jara, A. J., & Fernández-Breis, J. T. *From orbit to curb: Remote sensing and data-driven AI for public-transport observability, traffic, and parking in smart and sustainable cities: A critical survey*. MDPI Sustainability, 2026.

Constellr, *Technology*. https://www.constellr.com/technology

Cuevas Martínez, I., Jara, A. J., & Fernández-Breis, J. T. *Commercial Earth observation for wildfire early warning: A FinOps framework for data spaces*. IEEE, 2026.

Drouyer, S., & de Franchis, C. (2020). Parking occupancy estimation on Sentinel-1 images. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, V-2-2020*, 821–828. https://doi.org/10.5194/isprs-annals-V-2-2020-821-2020

Encío, L., Díaz, C., del-Blanco, C. R., Jaureguizar, F., & García, N. (2023). Visual parking occupancy detection using extended contextual image information via a multi-branch output ConvNeXt network. *Sensors, 23*(6), 3329. https://doi.org/10.3390/s23063329

EOI Space. *Stingray satellite imagery*. Retrieved September 13, 2026, from https://eoi.space/satellite/

European Union Aviation Safety Agency. *Operating a drone*. Retrieved September 13, 2026, from https://www.easa.europa.eu/en/domains/drones-air-mobility/operating-drone

Guth, A., Dauner, M., Coffey, E. R., & Hannigan, M. (2026). Using low-cost sensors for fenceline monitoring to measure emissions from prescribed fires. *Sensors, 26*(2), 745. https://doi.org/10.3390/s26020745

Haddad, T. (2026, January 23). *Clarity-1: What worked, and where we go next*. Albedo. https://albedo.com/post/clarity-1-what-worked-and-where-we-go-next

ICEYE. *Imaging modes*. Retrieved September 13, 2026, from https://www.iceye.com/defense-and-intelligence/imaging-modes

Kim, D., Yu, J., Yoon, J., Jeon, S., & Son, S. (2021). Comparison of accuracy of surface temperature images from unmanned aerial vehicle and satellite for precise thermal environment monitoring of urban parks using in situ data. *Remote Sensing, 13*(10), 1977. https://doi.org/10.3390/rs13101977

Lou, L., Zhang, J., Xiong, Y., & Jin, Y. (2019). An improved roadside parking space occupancy detection method based on magnetic sensors and wireless signal strength. *Sensors, 19*(10), 2348. https://doi.org/10.3390/s19102348

NASA. *Landsat 8*. Retrieved September 13, 2026, from https://science.nasa.gov/mission/landsat-8/

NASA Earthdata. *VIIRS I-band 375 m active fire data*. Retrieved September 13, 2026, from https://www.earthdata.nasa.gov/data/instruments/viirs/near-real-time-data/viirs-i-band-375-m-active-fire-data

Peraza-Garzón, J., Huerta-Mora, E., Olivarría-González, M., Quiñonez, Y., Rubio-Ayala, H., Palacios-Navidad, J. A., & Peraza-Garzón, A. (2026). Intelligent car park occupancy monitoring system based on parking slot and vehicle detection using DJI Mini 3 aerial imagery and YOLOv11. *AI, 7*(2), 74. https://doi.org/10.3390/ai7020074

Pereira, G. H. de A., Fusioka, A. M., Nassu, B. T., & Minetto, R. (2021). Active fire detection in Landsat-8 imagery: A large-scale dataset and a deep-learning study. *ISPRS Journal of Photogrammetry and Remote Sensing, 178*, 171–186. https://doi.org/10.1016/j.isprsjprs.2021.06.002

Pesonen, J., Hakala, T., Karjalainen, V., Koivumäki, N., Markelin, L., Raita-Hakola, A.-M., Suomalainen, J., Pölönen, I., & Honkavaara, E. (2024). *Detecting wildfires on UAVs with real-time segmentation trained by larger teacher models* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2408.10843

SatVu. *Thermal imagery*. Retrieved September 13, 2026, from https://www.satellitevu.com/products/thermal-imagery

Sheehan, A., Beddows, A., Green, D. C., & Beevers, S. (2023). City scale traffic monitoring using WorldView satellite imagery and deep learning: A case study of Barcelona. *Remote Sensing, 15*(24), 5709. https://doi.org/10.3390/rs15245709

STAC Community. *The STAC specification*. Retrieved September 13, 2026, from https://stacspec.org/en/about/stac-spec/

Stopic, R., Dias, E., de Kleijn, M., & Koomen, E. (2023). Satellite parking: A new method for measuring parking occupancy. *AGILE GIScience Series, 4*, 44. https://doi.org/10.5194/agile-giss-4-44-2023

TROPOMI. *Nitrogen dioxide.* Retrieved September 13, 2026, from https://www.tropomi.eu/data-products/nitrogen-dioxide

UP42. *BlackSky*. Retrieved September 13, 2026, from https://docs.up42.com/data/blacksky

Voogt, J. A., & Oke, T. R. (2003). Thermal remote sensing of urban climates. *Remote Sensing of Environment, 86*(3), 370–384. https://doi.org/10.1016/S0034-4257(03)00079-8

Wei, J., Li, Z., Lyapustin, A., Wang, J., Dubovik, O., Schwartz, J., Sun, L., Li, C., Liu, S., & Zhu, T. (2023). First close insight into global daily gapless 1 km PM2.5 pollution, variability, and health impact. *Nature Communications, 14*, 8349. https://doi.org/10.1038/s41467-023-43862-3

Xu, W., Wooster, M. J., He, J., Meraner, A., Gomez-Dans, J., Liu, Z., Trigo, I. F., & Dutra, E. (2026). Major improvements in spaceborne early fire detection and small-fire FRP retrieval with the Meteosat Third Generation Flexible Combined Imager. *Science of Remote Sensing, 13*, 100366. https://doi.org/10.1016/j.srs.2026.100366