World Journal of Advanced Engineering Technology and Sciences
eISSN: 2582-8266
Cross Ref DOI: 10.30574/wjaets
Journal homepage: https://wjaets.com/

(RESEARCH ARTICLE)

Check for updates

SUSTAINABLE AI DATA CENTERS: A DATA-DRIVEN FRAMEWORK FOR PREDICTING NATURAL RESOURCE STRESS, PUBLIC HEALTH IMPACTS, AND NEXT-GENERATION INFRASTRUCTURE SOLUTIONS

Hema Ravi Teja Bollam *

JPMC, Consultant - Lead Software Engineer.
* Corresponding Author

World Journal of Advanced Engineering Technology and Sciences, 2026, 21(02), 100–109

Publication history: Received on 06 September 2026; revised on 02 November 2026; accepted on 04 November 2026

Article DOI: https://doi.org/10.30574/wjaets.2026.21.2.0472

Copyright © 2026 Author(s) retain the copyright of this article. This article is published under the terms of the Creative Commons Attribution License 4.0

ABSTRACT

Whether replying to a prompt, creating an image or an update to a recommendation, each reply requires the use of a physical data center that consumes electricity and, in most instances, requires water cooling. This dependence has increased at a rate faster than the planning tools designed to monitor this. Global demand for electricity by data centres is projected to grow by over twofold from 2024 to 2030, and the amount of water that many draw from local rivers, aquifers and electric utility systems is growing at an even quicker rate in actual terms – with the demand directed to a small number of grid areas and watersheds that were not designed to handle it. Meanwhile, the back-up generators that keep campuses open during times of grid stress or while maintenance is being performed have led to complaints from residents in Loudoun County, VA to Boxtown, Memphis to Milwaukee County, WI about air quality. The three problems – pressure on the electrical grid, pressure on water supplies, exposure of the community for health problems and the selection of cooling infrastructure – are being addressed by different departments at different times with different data sometimes several years apart. This paper presents a proposed approach that unites all three in a single shared regional data layer, and a single health-exposure estimate and a single infrastructure recommendation, both produced in plain language, for any single given facility. We outline the data layer, the three prediction models derived from it and provide a detailed example of identifying a typical case many months before a permitting dispute arose. We conclude with a description of the way in which the approach might be evaluated or tested, some thoughts about how it might be put into practice by utilities, regulators and operators, and an explanation of when and when not this type of forecasting is useful.

Keywords: AI Data Centers; Electricity Grid Stress; Water Resource Stress; Public Health Exposure Assessment; Air Quality and Generator Emissions; Data Center Siting And Permitting; Liquid Cooling (Direct-To-Chip and Immersion Cooling); Regional Data Integration; Environmental Justice; Sustainable Infrastructure Planning; Predictive Resource Modeling.

1 INTRODUCTION

In February 2026, residents of the state of Virginia's Loudoun County were informed that state regulators are considering allowing data centres to keep their backup diesel generators on for longer and more often during the hottest part of the year; Loudoun County is home to thousands of data centres [1] [2]. Meanwhile, a supercomputer complex

was facing scrutiny in the Boxtown section of Memphis, some 800 miles south, for its lack of adequate permitting or pollution controls by community groups, which are concerned that it was operating dozens of gas turbines [3]. A health study conducted by Milwaukee County gave monetary estimates of the health harm that would happen to nearby residents every year due to their use of backup generators at a proposed campus [4]. None of these cases on their own suggests that there is a data issue—we have seen a rapid growth in computing power over the years that makes a facility necessary—and these facilities are designed to service the new power requirements, but their impacts on the water table as well as on air quality and household health are only felt later, and only by residents and journalists, not necessarily foreseen at the outset.

The energy use is no longer just a small part of the country's total energy use. The International Energy Agency's most recent figures show that global electricity consumption by data centres was approximately 415 terawatt-hours in 2024 and is expected to increase to 950 terawatt-hours by 2030—the rise being mainly due to data centres involved in AI, whose usage is set to grow from about 103 terawatt-hours to 465 terawatt-hours during that time [5]. Likewise, water usage in data centres is predicted to more than double by 2030, reaching approximately 9.3 trillion litres per year, and the land area used will more than double from around 6,900 to over 14,500 square km, according to a United Nations University Institute for Water, Environment and Health 2026 assessment [6]. While these are global numbers, the impact is not felt equally across all of the world—it is experienced in individual counties, along certain rivers, and by individual utilities; some are already urging hyperscale operators to either slow their operations or relocate their planned campuses.

The critical thing about this manuscript is that resource strain, impact on community health and decisions about infrastructure are not three distinct issues that can all be found within the same structure. When that's the case, they actually are three different views of exactly the same selection—where to locate the facility, how to power it, and how to cool it—and should, as a result, be forecasted by all the same data, not by three different departments on three different schedules. We present a framework based on a single regional data layer which provides three linked prediction models, as explained in Section 4, and we look at a typical case in Section 6 to show the practical differences that result from such integration.

2 THE CURRENT PICTURE

For the AI campus, the actual needs of the grid and the watershed are analyzed and listed.The actual needs of the grid and the watershed for the AI campus are analysed and listed.

The clusters of AI training or inference currently deployed are packed so tightly that they were not as densely-packed as older enterprise data centres. Today, the power consumption of a rack is around 10 kW on average, and next-generation accelerator racks are anticipated to consume 600 kW by 2027 [7]. In all of this power, it has to be extracted from the building, and for the overwhelming majority of facilities today (70–80 per cent) the most economical method of extracting the heat on a large scale is evaporative cooling, which is the process of using water to remove heat in exchange for electricity [8]. This exchange has been estimated to be responsible for the use of approximately 17 billion gallons of water directly by data centres in 2023 and an additional 211 billion gallons of water indirectly through the power plants that provide electricity to the data centres [7]. For example, in Texas alone, 25 billion gallons was consumed in 2025, and the regional AI researchers predict that consumption could rise to 161 billion gallons per year by 2030 as AI infrastructure grows in the state [8].

The following two resource curves (electricity and water) show how the two resources are projected to evolve over the next decade (Figure 1 and Figure 2). The separation in Figure 1 between lines for AI and for all data centres is part of this story that did not exist 10 years ago – it's an indicator of new demand being imposed on top of infrastructure that was already stretched or strained, in many cases.

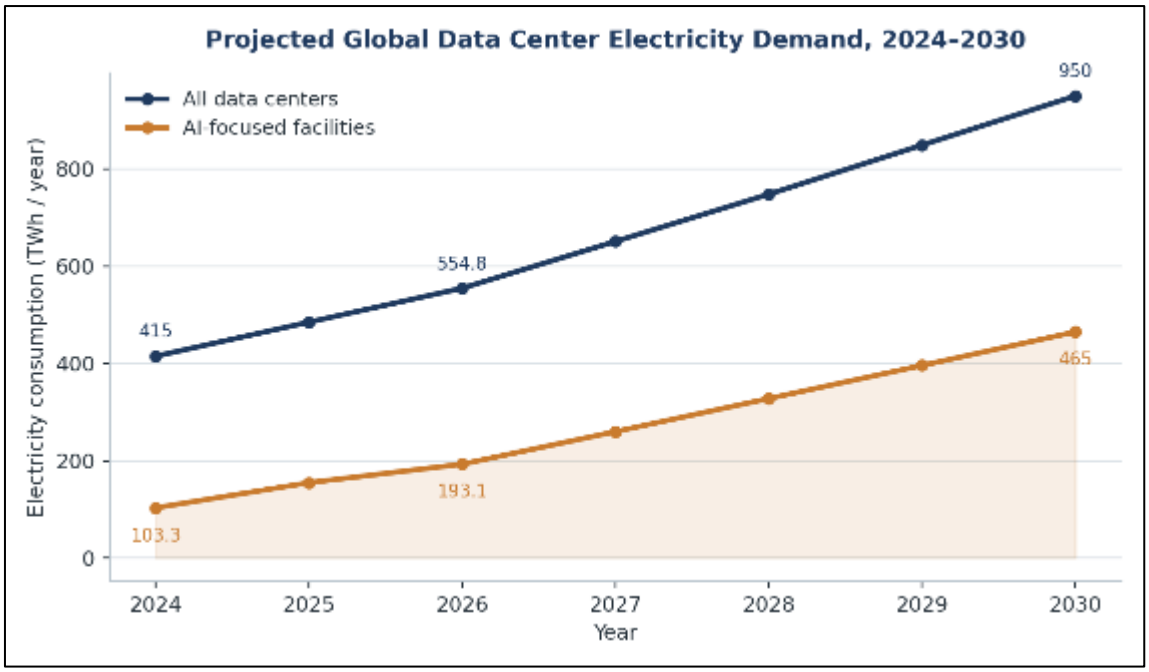


Figure 1: A comparison of the projected global electricity demand from all data centres with that from data centres focused on AI, from 2024 to 2030. The anchor years (2024, 2025, 2026 and 2030) are according to IEA estimates, the years in between having been interpolated. Source: [1].

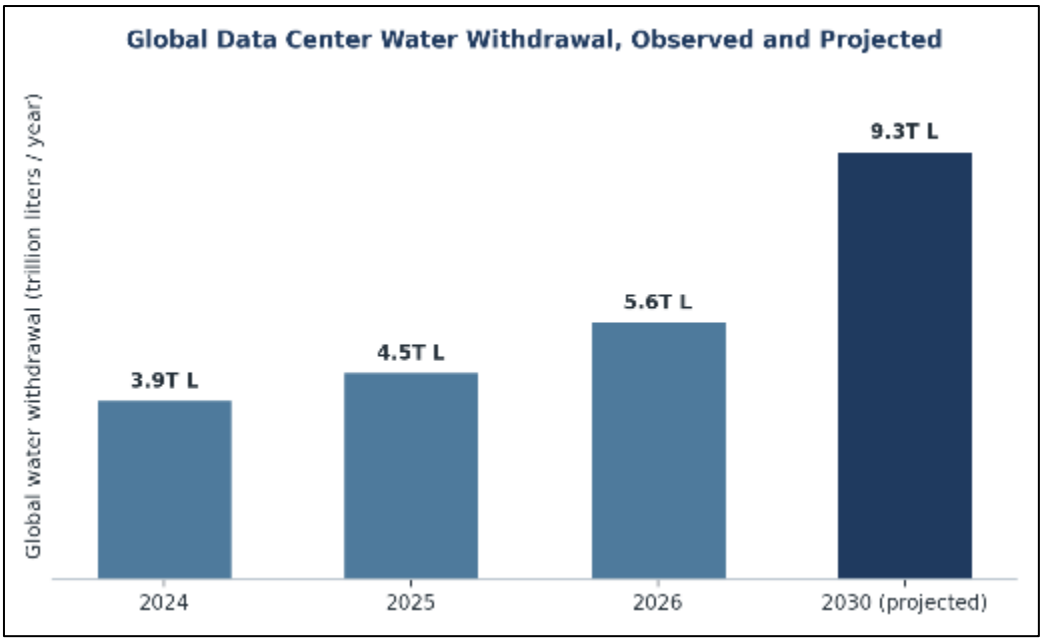


Figure 2: Global data center water withdrawal, observed and projected

An estimate of the global water withdrawal by data centres between 2024 and 2030 is given in Figure 2; the 2026 figure is an estimate for the present year and not a complete annual total. Sources: [1], [2], [3].

The trend is also illustrated in Table 1 when the land footprint is taken into account, since the locations of these campuses — in other words, the places where they are physically situated — have now become a resource constraint in regions where suitable, grid-connected and water-accessible land is scarce.

Table 1 Trend

Year	Electricity — all data centers (TWh)	Electricity — AI-focused (TWh)	Global water withdrawal (trillion L)	Land footprint (km^2)
2024 (observed)	415	103.3	≈3.9	≈6,900
2025 (observed)	485	155	≈4.5	—
2026 (year-to-date estimate)	554.8	193.1	≈5.6 (est.)	—
2030 (projected)	950	465	9.3	>14,500

The data on global electricity demand, water withdrawal, and the land footprint of data centres between 2024 and 2030 are given in Table 1; the figures for 2026 are preliminary annual estimates and not final totals; furthermore, the cross-source estimates of water vary slightly according to the method employed. Sources: [5], [6], [9], [10].

2.1 What is observed at the neighborhood level

The health impact comes one level below the overall statistics, especially in the counties where these facilities are found. Backup power is the clearest example of this. In just Northern Virginia there are over 10,500 generators connected to data centres, the vast majority of them being diesel-powered [1]. A study carried out by researchers from Virginia Commonwealth University and the University of Washington, as reported by the Washington Post, found that even if these generators only run for less than an hour each week on average, the effect on public health is the same as that of five power stations the size of the gas-fuelled Possum Point station near Quantico operating continuously — a level which, according to that study, is sufficient to worsen lung and heart conditions in the area and to result in at least three premature deaths each year [11]. A separate study by researchers at the University of California, Riverside and Caltech estimated that the nitrogen oxides and fine particulate matter from diesel exhaust in that region cause about 14,000 asthma symptom episodes every year [11].

Memphis is another example of this problem. Community groups have noted that 35 gas turbines at a supercomputer facility in the Boxtown area have no proper pollution controls or the required permits, a situation which has drawn the attention of both state and federal authorities because of nitrogen oxide emissions in a community that is already bearing an unfair share of pollution [3]. In Milwaukee County, an air-quality study carried out at the county's request estimated the local health cost at about $100,000 each year in that county alone and nearly $1 million each year in the wider region, on the basis that the backup generators operate at their full permitted level [4]. On a national scale, the AI Now Institute's data centre tracker predicts that air pollution resulting from the growth of data centres could cause an additional 600,000 asthma cases and about 1,300 premature deaths each year in the United States by 2030 if current backup-power practices continue unchanged [3]. Table 2 sets these documented cases out side by side; they come from different studies using different methods, so the figures cannot be simply added together, but the overall pattern among them is consistent.

Table 2: The community health effects that have been recorded or estimated in relation to data centre backup power in the cases reported for 2026. The figures come from separate studies employing different methods and so cannot be summed up across the rows. Sources: [7], [8], [9].

Location	Observed condition	Estimated health burden	Source
Northern Virginia (Loudoun, Prince William, Fairfax)	10,500+ backup generators, mostly diesel	Comparable to five continuously running gas power plants; ≥3 premature deaths/year regionally	VCU / Univ. of Washington analysis, via Washington Post [12]
Northern Virginia (same region)	Diesel exhaust: NOx and fine particulates	≈14,000 asthma symptom episodes/year	UC Riverside & Caltech study [8]
Memphis, TN (Boxtown neighborhood)	35 gas turbines, limited pollution controls	Elevated NOx exposure in an already burdened neighborhood	AI Now Institute tracker [7]
Milwaukee County, WI	Proposed campus, full-allowance generator use	≈$100,000/year county-level; ≈$1M/year regional health cost	County analysis via Spectrum News 1 [9]
United States (national, 2030 projection)	Cumulative data-center-linked air pollution	≈600,000 additional asthma cases/year; ≈1,300 premature deaths/year	AI Now Institute [7]

2.2 In situations where the industry is already making a response

This all occurs within a wider context. Because of delays in getting water permits and resistance from the local community, operators are now using cooling systems that need far less water per h unit of heat removed. Both direct-to-chip cooling—this means passing a coolant through cold plates that are attached to the processors—and full immersion cooling—where the equipment is placed into a dielectric fluid—greatly cut down on evaporative water loss and make it possible to use the higher rack densities that are required by AI hardware [13], [14]. Data centres that use liquid cooling have achieved power usage effectiveness (PUE) ratios as low as 1.05 to 1.15, unlike the range of 1.4 to 1.8 seen with conventional air-cooled data centres, and companies have reported water consumption reductions of up to 50 per cent in instances where liquid cooling has completely taken the place of evaporative cooling towers [13-16]. The trend is towards this kind of cooling: spending on liquid and immersion cooling equipment is expected to reach about $1.1 billion in 2026, accounting for around 19 per cent of all spending on thermal management, with the immersion portion alone growing at nearly 25 per cent per year [13], [14]. Regulation is also contributing to this shift—Germany's Energy Efficiency Act will require new data centres to have a PUE of 1.3 or better by 2027, and the EU's broader Energy Efficiency Directive sets similar targets across the region [13].

2.3 Finland waste-heat recovery

This placement establishes heat recovery as an existing response before Section 4 converts it into a framework recommendation. Helsinki has reused data-centre heat through its district-heating network since 2010, and subsequent projects have extended that approach to additional sites and buildings [17], [18].The technology required to reduce both water use and, indirectly, the need for diesel backup is already available and is being developed. The one thing that is missing is a way of beforehand and for a specific proposed site assessing whether the local power grid and watershed are able to cope with a given facility, what effect the facility's backup power plan would have on the surrounding area, and which cooling design in fact takes into account the particular constraints of that location—rather than having to discover the answers during the permitting process after the site has already been chosen.

2.4 Problem Statement

This paper suggests a framework which suffers from three structural defects. The first of these is that the right decisions are made in the wrong order. Site selection, power procurement, and the acquisition of permits for backup generators are typically carried out first, driven by the cost of the land, access to fibre, and one's position in the grid interconnection queues; the water utility, the air-quality regulator, and the county health department only get involved later, generally after construction has already begun, at a stage when the only possible alternatives left are measures designed to reduce the impact rather than altering the design. The second defect is that the data needed to forecast resource demand and health exposure is already available but is stored in separate systems and is seldom integrated: the load data from the grid balancing area is kept by the transmission operator, the permits for water withdrawal are held by the state environmental agency, information on generator run time and emissions can be found in the air permits filings, and records relating to asthma and cardiovascular health are maintained by the county public health departments. None of these individual datasets contains the information that a person making a siting decision requires. The third problem is that, in this field, whenever models are used they are generally of a single kind — for instance, a study on grid interconnection, a study on water availability, or an air dispersion model — and are commissioned separately, at different times, by different consultants, with no common assumptions and no way in which the results of one model can be used to update another.

Recent disputes show how these structural defects can also develop into conflict over permitting. In Monroe Township, New Jersey, the municipal council adopted an ordinance prohibiting data centres after residents opposed a proposed project, and Hexa Builders later filed a lawsuit seeking $300 million in damages [19], [20]. The developer maintains that the project had received preliminary approval and that standards concerning noise, vibration, water use, and energy use were being discussed before the ban; these remain claims made in an unresolved dispute rather than established findings [20]. The case nevertheless shows the practical cost of allowing community concerns, technical safeguards, and land-use decisions to proceed on separate tracks until the permitting process becomes adversarial. A shared pre-siting assessment would not necessarily remove disagreement, but it could give residents, planners, and the developer a common evidence base before major commitments are made.

A good framework has to meet four aims that current practices do not: it ought to give a forward-looking and continuously updated forecast of resource strain rather than depending on a single study; it should connect this forecast to a specific and quantified evaluation of health exposure in the nearby area rather than treating air permitting as a separate matter; it should propose an infrastructure response—for example, cooling technology, a modified backup-power setup, or a change in siting—one that is appropriate for the real resource and health constraints present at that location; and it must present all of this in a form that a county planning board, a state utility commission, or a community advisory group will actually be able to read and question, rather than issuing a technical report that only the consultants who put it together can understand.

3 THE FRAMEWORK

The system makes use of one regional data layer to provide three linked prediction models, and on this is situated a simple report which enables the output to be made available to those people who are responsible for making decisions regarding siting, permitting and design in a form that they can actually use.

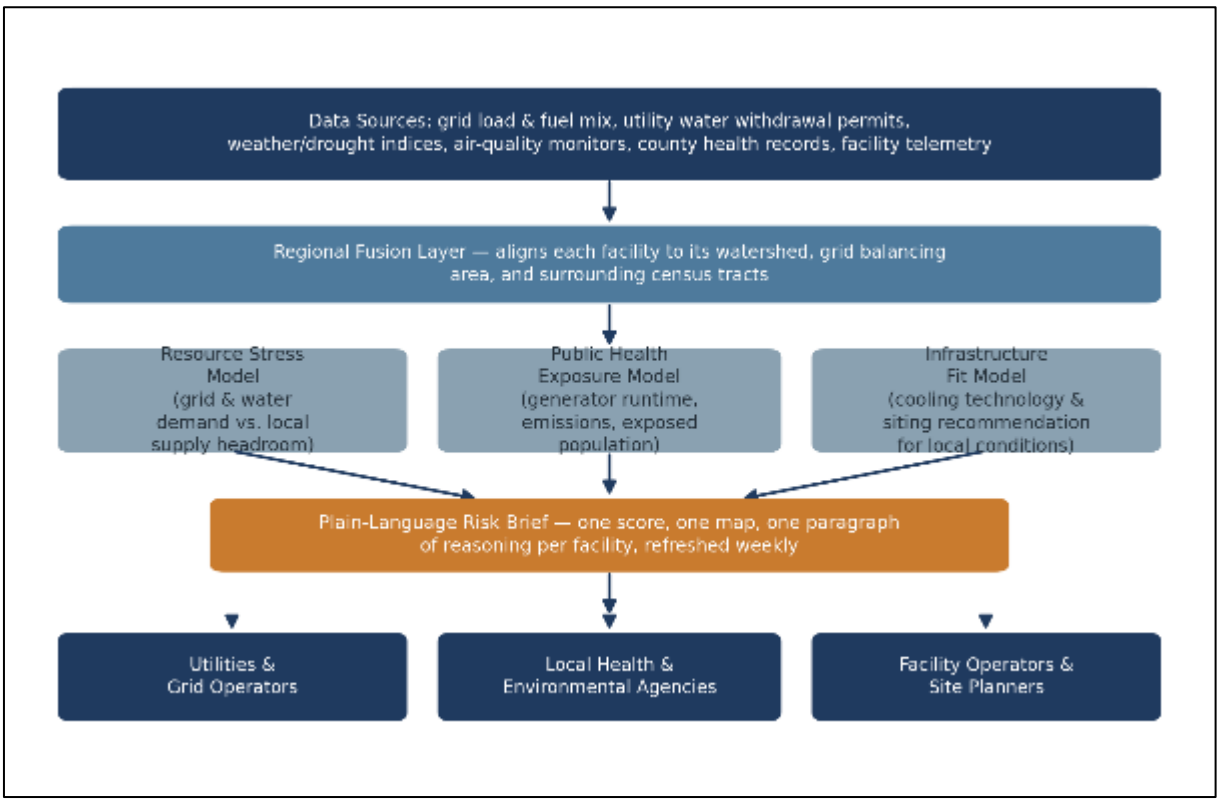


Figure 3: The architecture that is proposed is illustrated in Figure 3 and includes a common regional data layer that feeds three interconnected prediction models—models concerned with resource stress, public health exposure, and infrastructure fit—and these models combine to generate a single summary in plain language which is given to utilities, health and environmental agencies, and facility planners.

3.1 The layer concerned with regional data

Each input is linked to a specific facility and, most important of all, to the particular watershed, grid balancing area, and set of census tracts that surround it—rather than to a national or state average. The inputs include data on grid load and the fuel mix used in generation obtained from the relevant balancing authority; information on water withdrawals together with drought or reservoir level indices from the state environmental agency and the regional water utility; readings from the air-quality monitors and records of generator emissions; county-level data on cases of asthma, visits to emergency departments, and cardiovascular admissions, all at an aggregation level that protects individual privacy; and, where available, telemetry from the facility relating to the actual power drawn, the type of cooling system, and the generator run time. At this stage the tedious but essential work is carried out of bringing all this data into line with a common geographical and temporal framework, so that a rise in generator run time reported in the permit filings can be compared with the air-quality readings from the same week and the emergency department data from the same month for the surrounding tracts.

3.2 Resource stress model

Rather than compare a facility's anticipated electricity and water demand with the actual remaining headroom in its local grid balancing area and watershed—rather than using national averages, which generally underestimate the strain in the specific counties where the campuses are located—it generates a stress score that is continuously updated as new grid and drought data come in, taking into account both the present situation and the rate at which the local headroom is being consumed by the facility itself and by other nearby projects that are using the same grid segment or the same water source. The increasing stress score is the very first and least controversial signal that the framework provides, since it is based on supply-and-demand calculations that utilities and regulators already keep an eye on—the innovation here is that it makes this type of accounting visible in a forward-looking manner for each facility and each region, instead of having to reconstruct it after a shortage or curtailment has occurred.

3.3 The public health exposure model

The model takes the generator and turbine operating times — which come from permit applications and, if possible, from actual telemetry data — applies the emission factors that are known for the relevant type of fuel and equipment class, considers the local wind and dispersion patterns, and then overlays the resulting exposure area against the population that actually resides within it. The outcome is an estimate of the excess health impact — a kind of figure akin to those for Virginia and Milwaukee that are summed up in Table 2 — and this estimate is provided specifically for a certain proposed or currently operating facility rather than as an average for the whole country or the entire industry. As the case studies in Section 2.2 show that these effects can be significant even when generator use is low, the model is deliberately designed to assess risk at the maximum generator operating time permitted by the permit, not just at the level of historical usage, since permits are usually set to allow a great deal more generator operation than facilities actually carry out in a typical year.

3.4 Infrastructure fit model

Starting from a resource-stress profile and a health-exposure evaluation for a specific location, the model recommends a cooling and backup-power arrangement that is suitable given that location's particular situation, rather than merely going with the type of setup that a particular vendor or operator normally uses. Table 3 shows the trade-offs that the model takes into account. For instance, in a area where water is scarce but there is sufficient capacity on the grid, it would favour direct-to-chip or immersion cooling together with grid-interactive battery backup instead of relying on diesel generators. Conversely, in a region where the grid is limited but water is available, it might instead opt for carrying out efficiency retrofits on the existing air-cooled system along with demand-response agreements. The important thing is not that liquid cooling or batteries are always the best choice — it is that the right option depends on the local conditions, which the current siting process very infrequently quantifies before construction begins.

Heat recovery provides a further response that is especially relevant in cold climates. In Helsinki, an underground data centre built in a former civil-defence shelter beneath Uspenski Cathedral was connected to the city’s district-heating system so that heat produced by the servers could be recovered and used in nearby buildings. The city’s energy company has since continued to integrate data-centre waste heat into its district-heating network [17], [18]. This example does not mean that underground construction or heat recovery is appropriate at every site. It shows, however, that where a district-heating network, suitable heat-pump infrastructure, and a nearby demand for heat are already present, the thermal output of a data centre can be treated as a useful local resource as well as a cooling load.

Table 3.The trade-offs linked with cooling technology as evaluated by the infrastructure fit model are given in Table 3; the figures for PUE and water use refer to ranges reported by the industry and are not based on measurements obtained at any specific facility. Sources: [13], [14], [15].

Cooling approach	Typical PUE	Relative water use	2026 deployment stage	Best fit
Air-cooled (CRAC/CRAH + cooling towers)	1.4 – 1.8	Baseline (highest)	Legacy majority of installed base	Low-density legacy halls; water-rich sites only
Direct-to-chip (cold plate, single/two-phase)	1.1 – 1.3	30–50% below air-cooled baseline	Scaling quickly; ≈19% of 2026 thermal-management spend	High-density AI racks (60–120+ kW); retrofits
Full immersion cooling	1.05 – 1.15	Near-zero evaporative loss	Early but growing at ≈25% CAGR	New-build hyperscale campuses in water-stressed regions

3.5 Shared reporting layer

Each brief includes for each facility a single stress score, illustrating the area surrounding the facility to which there is exposure, and a short paragraph explaining which inputs were used to arrive at the score and why it was reached. The brief is updated on a fixed schedule—weekly is a reasonable default—rather than being prepared once at the permitting stage and then left to become out of date. The objective is that someone who is a member of a county planning board or a member of a community advisory group should be able to read the brief without possessing any technical knowledge and then be able to ask a specific and well-informed question about it, which is a lower standard than being able to understand the underlying grid-load or air-dispersion model directly, and a much higher standard than is achievable with most present environmental review documents.

4 A REPRESENTATIVE SCENARIO

Let us consider a situation following the pattern typical of those described in Section 2.2, rather than being based on any specific facility. A large-scale operator is proposing an AI training campus with a capacity of 150 megawatts in a mid-Atlantic county that already hosts several data centres on the same grid segment and making use of the same regional aquifer. The site plan includes conventional evaporative cooling and also has a set of Tier II diesel generators capable of handling the facility's entire emergency load, in line with the standard practices in the area.

When considering the proposal, the resource stress model shows that the local grid section is already operating with a very small safety margin as a result of the nearby facilities, and that the aquifer supplying the site's water utility has seen a decline in water levels over the past two summers—this situation alone could provide a reason to request a phased interconnection or an alternative water source, but would not by itself stop the project from proceeding. If the public health exposure model is applied to the generator bank's permitted maximum run time rather than to its expected normal level of use, it calculates an exposure area that encompasses two residential subdivisions and an elementary school within a distance of about two miles, leading to a health burden estimate similar to the one given by Milwaukee County in Table 2. Taken together, these two findings would cause the facility's overall assessment to be placed in a review category which requires the preparation of a documented mitigation plan before permission can be granted—this does not amount to an automatic rejection, but instead involves a specific, evidence-based discussion that takes place before groundbreakings take place rather than after residents have started to make complaints.

In this situation the infrastructure modifications recommended by the model are a direct consequence of the diagnosis: as there is a certain amount of grid headroom although it is only a small amount and the water headroom is more obviously limited, it is suggested to use direct-to-chip cooling rather than evaporative towers in order to reduce water withdrawal by an estimated 30 to 50 per cent according to the industry figures given in Table 3, and the battery storage should be sized to cope with routine outages so that the diesel generators are only used in actual prolonged emergencies and not operated at their full permitted level. None of these specific recommendations is in any way new since utilities, water authorities and cooling vendors make them separately during individual consultations. The difference is that they are now all brought together and supported by the same quantified evidence at the stage when the site plan still has room to be altered rather than after it has been finalised.

5 EVALUATION APPROACH

The evaluation should be carried out in two stages since this manuscript presents a design framework rather than the results of a real-world deployment. In the first stage, backtesting should be done by applying the resource stress and health exposure models to the cases described in Table 2, using only the data that would have been available at each facility when it was proposed, and checking that the models would have picked up the same problems that were later identified by researchers, journalists, and residents. This is a reasonable short-term way of validating the approach because the cases in question are already publicly available and therefore provide a natural benchmark before any pilot deployment takes place. The second stage is prospective and consists of implementing the framework in connection with a small number of new siting reviews, observing whether the predicted stress and exposure scores match what the utilities, the air monitors, and the health records show after construction has taken place, and then using the difference between the forecasts and the actual results to recalibrate the underlying models.

The useful metrics differ from one model to another. In the case of the resource stress model, it is important to examine how far in advance it detects a grid or water shortage that later leads to the need for curtailment, an increase in rates, or an emergency limitation of water use, as well as how often it sends out a warning which turns out to be unnecessary—since if the number of false alarms is too high, the planning boards will ignore the tool even though it may be accurate in all other respects. As for the health exposure model, it should be compared with the real emergency department and respiratory complaint data from the years following the facility was constructed, bearing in mind that there are numerous factors which affect health outcomes and that the model's estimate should be seen as just one of several inputs rather than as a definite medical judgment regarding any individual. The infrastructure fit model should be assessed by looking at both adoption and outcomes: that is, whether the operators who receive a recommendation to cool actually implement something that is close to the recommendation, and whether the facilities that actually show the reduced water use and lower emissions which the model had predicted do so in fact.

6 FACTORS TO TAKE INTO ACCOUNT CONCERNING DEPLOYMENT AND GOVERNANCE

Only when forecasting is done by an organisation which the community in question already trusts — for instance, a state environmental agency, a public utility commission, or a university research partnership — or in cooperation with such an organisation will it be able to win the community's confidence; it would not be acceptable for the facility operator to prepare the assessment on their own. Because health data is the most sensitive type of input in the

framework, it should be used only at a geographical level of aggregation which cannot be linked to individuals, in keeping with the usual practices for sharing public health data; the exposure model needs information on population density and general health-trend data for a census tract, not individual medical records. The various scores generated by the framework must be reproducible on the basis of the data it references, so that a community group or a consultant who holds opposing views can check the framework's calculations rather than simply being asked to accept the figures. Finally, since the framework is specifically intended to help with decisions before a facility is constructed, it must include a way of feeding into the actual permitting process — a resource-stress score that is produced after a facility has already been built and is operating may be of diagnostic value but has mostly lost the ability to affect the outcome.

6.1 Limitations

The data given in Section 2 are drawn from a wide range of sources – such as government, academic, and industrial ones – each of which uses different boundaries and methodologies, and hence the figures must be seen as the best current estimates rather than as precise measurements; specifically, the figures for water vary considerably depending on whether or not the indirect water use linked to electricity generation is included. The health burden estimates in Table 2 are based on a small number of published studies that focus on particular areas of the United States; in order to apply this kind of exposure modelling to a new location it is necessary to have local emission, dispersion, and population data, which may not be easily available in all places, and the dose-response relationships between generator emissions and specific health effects have their own scientific uncertainties, a point which this framework does not deal with. Furthermore, the framework itself is a design proposal and not a validated system: the backtesting and prospective evaluation referred to in Section 6 have not yet been carried out, and the case-study walkthrough in Section 5 only shows the logical structure of the framework and does not describe a real-world implementation. Finally, no forecasting tool has the ability by itself to change a siting decision; it can only affect what decision-makers know at the moment when the decision is being made, and consequently it is still up to policy and the public process as to what they do with that information, not something that the model can decide.

7 CONCLUSION AND FUTURE RESEARCH

We already possess the data needed to judge, long before any major development takes place, whether a proposed AI data center will place a strain on the local power grid, lead to a drop in an already under aquifer, or expose a neighborhood to substantial pollution from backup generators. So far, however, there has been no single integrated model which combines information about the grid, water resources, permitting, and health and presents the results in a form that can be understood by those people who are responsible for deciding where such sites should be located and for granting permits, instead having three separate reports prepared by three different experts on different schedules and these rarely being compared with each other. The framework suggested here — made up of one regional data layer, three linked prediction models, and one short written explanation in plain language — is offered as a first step towards achieving this integration, drawing on situations that have already happened in Virginia, Memphis, and Wisconsin, and on cooling technology currently available on the market and which is being scaled up.

In order for the framework to move from being a design proposal to a tested tool, three steps are required. The first of these is to backtest the framework using a wider range of documented cases than those included in this report in order to examine how consistently its logic would have foreseen the actual disputes. The second step consists of carrying out a pilot deployment in conjunction with a state agency or a utility commission that is willing to apply the framework to new siting reviews, making use of the data derived from comparing the forecasts with the actual results to recalibrate the models. The third step is to extend the dose-response assumptions of the health exposure model in collaboration with public health researchers, since this component is based on the smallest number of well-established local relationships of the three models and would therefore benefit most from specific epidemiological input. None of these steps involve the need for new sensors or new scientific research; rather, they require the agencies which currently hold the data to make it available to a shared model before the next site plan is finalised rather than after.

REFERENCES

[1] Gradually.ai. (2026). How much electricity and water AI data centers are using at the moment.

[2] Reuters. (2026). UN researchers say AI will double data centre power and water use by 2030. Insurance Journal.

[3] Business Today. (2026). Why do AI data centers use so much water? Explained.

[4] Doty, T., Estrada Chavez, R., et al. (2026). AI data centers: Energy, water, and cooling demands. Boise State University Undergraduate Research Showcase.

[5] Borenstein, S. (2026). Energy, water use and pollution of AI and data centres rival most countries. Society of Environmental Journalists.

[6] Freshwater use in data centers. (2026). Research compilation.

[7] AI Now Institute. (2026). Air pollution. Data Center Tracker.

[8] National Community Land Conservancy Voters. (2026). Backup generators for data centers threaten community health.

[9] Spectrum News 1. (2026). Health impacts of data centres.

[10] Virginia Mercury. (2026). Lawmakers debate how to regulate data centers' diesel backup generators.

[11] PennEnvironment Research & Policy Center. (2026). PJM proposal to relax air pollution standards for data centers during extreme heat puts public health at risk [Statement].

[12] Piedmont Environmental Council. Take action: The air we breathe is at risk.

[13] GM Insights. Data center liquid cooling market: Size and forecast report.

[14] Cleantech Group. Revolutionizing data center cooling: Immersion technologies at the forefront.

[15] Barchart. Data center direct-to-chip coolants market outlook 2026: Growth drivers, key players, consumer trends, and forecast insights.

[16] Net Zero Insights. Solutions for cutting data center water consumption.

[17] Velkova, J. (2016). Data that warms: Waste heat, infrastructural convergence and the computation traffic commodity. Big Data & Society.

[18] Helen Ltd. (2022). Helen and Equinix extend successful cooperation to distribute more waste heat from data centers to properties in Helsinki.

[19] Township of Monroe, New Jersey. (2026). § 175-163.5 Data centers (Ordinance O:25-2026).

[20] Lorber, L. (2026). A New Jersey town banned data centers. One developer wants $300 million in damages. Inc.