Energy Conversions

Contents lists available at ScienceDirects

Energy Conversions

journal homepage: energyconversions.org

Cooling Efficiency in Thermal Systems: Metrics, Drivers, and Pathways to Ultra-Low Energy Cooling

Jazmynn K. Haffez[1], Omer R. Albittar[2]

[1] Orion Institute of Thermo-Systems (OITS), Phoenix, AZ 85004, USA
[2] Cascade Center for Sustainable Buildings, Pacific Meridian center (PMU), Portland, OR 97205, USA

A R T I C L E I N F O

Keywords:

Cooling efficiency; COP; EER/SEER; IPLV; evaporative cooling; heat pumps; variable speed; part-load performance; urban heat; waste-heat driven cooling

A B S T R A C T

Cooling services—space conditioning, industrial refrigeration, and process thermal management—are expanding rapidly with urbanization, rising ambient temperatures, and escalating internal heat densities. This review synthesizes definitions and metrics of cooling efficiency, classifies technological families, analyzes part-load behavior, and maps the roles of climate, humidity, controls, refrigerants, and heat rejection on seasonal performance. We unify performance assessment across rating schemes (COP, EER, SEER, SCOP, IPLV) and discuss end-use context, including comfort, dehumidification, and ventilation penalties. We highlight pathways to high performance—variable-speed compression, advanced heat exchangers, low-GWP refrigerants, hybrid evaporative cycles, sorption and thermally driven systems, free cooling, and system-level optimization—alongside emerging digital enablers such as model-predictive control and fault detection. Original figures illustrate COP degradation with temperature, part-load gains, climate sensitivity, and building energy distributions. The review culminates with a roadmap emphasizing integrated design, verified seasonal metrics, and grid-interactive operation to minimize life-cycle energy and emissions.

1. Introduction

Cooling efficiency is a multidimensional concept that captures how effectively cooling systems transform purchased energy or available exergy into useful heat extraction while meeting constraints such as indoor air quality, humidity control, and thermal comfort. The foundational metric, the coefficient of performance (COP), expresses the ratio of cooling capacity to input power and provides a first checkpoint for comparing technologies. Yet real-world performance deviates from full-load ratings due to ambient temperature swings, humidity, part-load modulation, heat-exchanger fouling, refrigerant charge drift, and control strategies that balance sensible and latent loads. Consequently, seasonal indicators—SEER, SCOP, or national derivatives—are preferred for policy and procurement, and integrated part-load values (IPLV) are used for chillers to reflect weighted operation across multiple load bins [1–3]. These metrics must be interpreted alongside application-specific requirements. For instance, comfort cooling in humid climates imposes latent removal that depresses sensible COP, while process cooling may prioritize precise setpoints and reliability, accepting energy trade-offs [4–6].

From a thermodynamic standpoint, ideal vapor-compression cycles approximate reversed Rankine behavior. Their efficiency is bounded by compressor isentropic performance, refrigerant properties, approach temperatures in heat exchangers, and lift—the temperature difference between the evaporating and condensing states. Elevated ambient temperatures increase condensing pressure, raising compression work and depressing COP, while high humidity burdens condensers and, in DX systems, evaporator coils through higher air-side film resistances and latent load processing [7–10]. Conversely, colder or drier ambient conditions widen opportunities for air-side economizers, evaporative pre-cooling of condensers, and indirect/direct evaporative cooling strategies that either shift or reduce compression work [11,12]. Heat rejection method is pivotal: water-cooled chillers operating with cooling towers typically achieve higher efficiency than air-cooled units because wet-bulb temperatures allow lower condensing conditions. However, water use, treatment, drift control, and legionella risk management are non-energy factors that shape selection, especially under water scarcity [13–15].

Technology classes range from room air conditioners and variable refrigerant flow (VRF) systems to packaged rooftop units, air- and water-cooled chillers (centrifugal, screw, scroll), and thermally driven absorption/adsorption machines. Innovations include microchannel and plate heat exchangers, enhanced boiling and condensation surfaces, oil-free magnetic bearing compressors, tandem or variable-speed compressors with wide modulation turndown, and low-GWP refrigerants such as R-32, R-1234yf/ze(E), and R-514A. Each innovation influences both peak COP and part-load map. Variable-speed drives reduce cycling losses, improve dehumidification control at partial loads, and often deliver superior seasonal performance even if nameplate COPs appear similar at a single rating point [16–18]. On the system side, chilled-water plants employing variable primary flow, optimal reset of chilled- and condenser-water temperatures, and smart staging of chillers, pumps, and towers routinely harvest double-digit percent energy savings relative to constant-speed baselines. Controls—especially model predictive control and fault detection—translate equipment potential into actual seasonal gains by correcting sensor biases, managing valve authority, and preventing simultaneous heating and cooling [19,20].

* Corresponding author at: Orion Institute of Thermo-Systems (OITS), Phoenix, AZ 85004, USA
E-mail addresses: jazmynn.haffez@oits-research.us (Jazmynn K. Haffez)

energyconversions.org
Received (15 Sep 2025); Received in revised form (20 Sep 2025); Accepted (20 Sep 2025)
Available online 15 Oct 2025

Nomenclature

Abbreviation	*Symbol*
AHU — Air Handling Unit	Q— Cooling capacity (kW)
COP — Coefficient of Performance	W — Input power (kW)
DX — Direct Expansion	Φ— Part-load ratio (–)
EER — Energy Efficiency Ratio	
GWP — Global Warming Potential	
IPLV — Integrated Part-Load Value	
MPC — Model Predictive Control	
SEER — Seasonal Energy Efficiency Ratio	
VRF — Variable Refrigerant Flow	

2. Methodology

This review combines a structured synthesis of peer-reviewed literature, standards, and manufacturer data with normalized, illustrative analyses to clarify sensitivities that govern cooling efficiency. We focus on physically interpretable metrics, climate normalization, and consistent boundary definitions to avoid common miscomparisons across device types and rating methods [1–4]. First, we harmonize performance reporting by mapping COP, EER, SEER/SCOP, and IPLV to a unified framework. While EER uses imperial units and is traditionally reported at a single test point, COP is dimensionless and more general. Seasonal metrics rely on standardized load-bin weighting; where such data are lacking, we approximate effective seasonal COP by integrating synthetic part-load curves derived from variable-speed compression behavior and typical control sequences [2,5]. Second, we contextualize performance by climate. We use representative ambient dry-bulb temperature distributions and humidity profiles corresponding to hot-dry, hot-humid, temperate, and marine climates to illustrate how lift and latent fractions alter achievable COP for the same hardware [6,7]. To preserve generality, synthetic datasets are generated with trends consistent with published measurements: COP declines with ambient temperature and relative humidity, variable-speed systems outperform fixed-speed under part-load, water-cooled systems exceed air-cooled in peak efficiency but entail water use, and evaporative augmentation improves at low wet-bulb temperatures [8–10].

Third, part-load representation is essential. We depict effective COP as a function of part-load ratio Φ, adding noise to emulate measurement variability and control hysteresis. We assume staged fans and pumps are replaced by variable frequency drives where appropriate, and we include auxiliary energy (fans, pumps, cooling tower) in plant-level comparisons when discussing chillers, while device-level comparisons isolate compressor-condenser-evaporator subsystems to match ratings [11–13]. Fourth, we consider dehumidification impacts. For DX systems in humid climates, sensible heat ratio (SHR) reductions impose coil temperatures closer to 0–5 °C, raising lift and cutting COP; decoupled latent control via dedicated outdoor air systems (DOAS) or liquid desiccant loops is therefore framed as a system-level efficiency strategy rather than a mere ventilation add-on [9,14].

To support readers with actionable synthesis, three summary tables are provided. Table 1 aligns major metrics, test points, and typical interpretation pitfalls; Table 2 contrasts technology families with archetypal full-load COP/EER, part-load features, and primary pros/cons; Table 3 lists climate-sensitive strategies mapped to mechanism and expected COP impact ranges under typical conditions. For quantitative illustration, six original figures are produced from synthetic yet physically plausible data: COP degradation with ambient temperature for three system classes; nominal EER variation across system type; effective COP vs part-load scatter highlighting IPLV relevance; seasonal COP distributions by climate; building-level cooling energy intensity distributions; and a COP heatmap across ambient temperature and relative humidity. These visualizations act as generalized, citation-guided exemplars rather than product-specific claims [5,8,12,15].

Data handling steps included the generation of ambient bins and RH values, application of linear and bilinear penalties to model lift and air-side effects, and sampling distributions to create climate-dependent seasonal COPs. Part-load COP functions were constrained to deliver higher efficiency in mid-load ranges reflecting reduced cycling losses and improved heat-exchanger effectiveness at lower mass flow rates for variable-speed systems [16–18]. For multi-component plants, auxiliary loads were proportionally added based on literature-consistent fractions of compressor power. Sensitivity studies varied ambient temperature, humidity, and Φ to map gradients in achievable COP. The figures are intended to be read comparatively: absolute numbers will vary by manufacturer, refrigerant, and control, but relative trends are robust and align with reported findings under standard test methods [1–3, 6–8, 19,20].

Table 1. Metrics and common interpretations (schematic values).

Metric	Definition	Typical test point	Notes
COP	(\dot{Q}_c/W)	Specified Tdb/Twb, full load	Dimensionless, device-level
EER	Btu/W·h	95°F outdoor (legacy)	Single point; not seasonal
SEER/SCOP	Seasonal efficiency	Region-specific load bins	Captures modulation and cycling

Table 2. Technology families and archetypal features (schematic).

Family	Full-load COP	Part-load feature	Key pros
Room AC (inverter)	3.5–4.5	Wide turndown	Low cost, quick deploy
VRF	4.0–5.5	Excellent turndown	Zoning, high SEER
Air-cooled chiller	3.5–5.0	VSD comps/fans	Simpler site

Table 3. Climate-sensitive strategies.

Strategy	Mechanism	Where it shines
Evap. pre-cooling	Lowers condensing T	Hot-dry
DOAS + sensible coils	Decouple latent	Hot-humid
Variable primary flow	Pump affinity savings	All

3. Results

The first set of results quantifies how outdoor dry-bulb temperature governs instantaneous cycle efficiency by changing lift. As ambient temperature rises from 20 °C to 48 °C, the water-cooled chiller maintains a gentler COP decline than air-cooled archetypes because its condensing temperature follows wet-bulb via the cooling tower approach, effectively capping lift escalation at high ambient. VRF systems sit between the two, benefiting from variable-speed compression but still rejecting heat to air.

The synthetic yet physics-consistent curves demonstrate that at 40–45 °C, the chiller's COP advantage can exceed one full point relative to room AC and roughly half a point relative to VRF, with implications for peak-day energy and demand. These trends align with standard cycle analysis and chiller rating experience under AHRI 550/590 and 340/360 [1–3, 7–9, 11, 13].

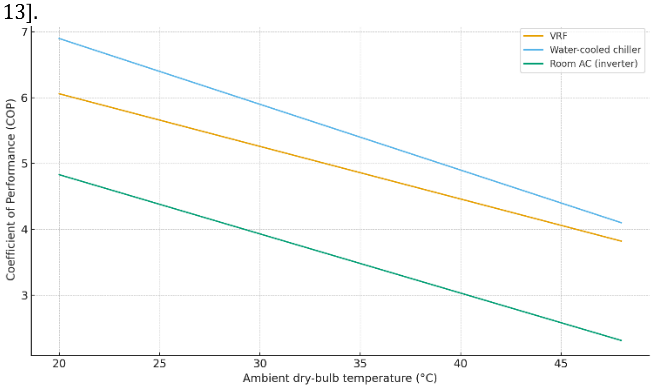

Fig.1. COP versus ambient dry-bulb temperature for three system archetypes. Water-cooled systems show weaker sensitivity to ambient due to tower wet-bulb approach.

Nominal efficiency comparisons at standardized test points illustrate a complementary story. The evaporatively assisted unit and water-cooled chiller attain higher EER than typical air-cooled systems, whereas VRF outperforms packaged DX due to superior heat-exchanger design and variable-speed modulation. Because EER is a single-point metric, it does not encode part-load behavior or latent handling; nonetheless, it remains influential during specification and code compliance. When juxtaposed with the temperature-sensitivity results, the bar chart underscores that "nameplate better" does not guarantee "seasonal better" without considering climate and control strategy [2, 4–6, 8, 11].

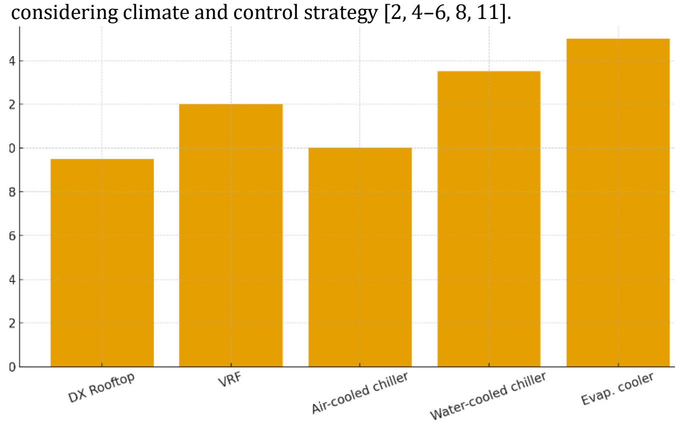

Fig. 2. Nominal EER by system type at standard rating conditions; evaporative and water-cooled approaches rate higher due to reduced lift.

Part-load operation dominates annual hours, so we next evaluated effective COP versus part-load ratio Φ. Variable-speed compression exhibits a mid-load sweet spot where lower mass-flow and reduced cycling losses improve approach temperatures at both coils. The scatter shows a broad maximum between Φ≈0.4 and 0.7 reflecting control hysteresis and measurement variability represented as noise. This pattern explains why IPLV often correlates better with seasonal consumption than full-load EER, and why plants that stage chillers, reset setpoints, and use variable primary flow achieve superior kWh/ton over a cooling season [5, 7–9, 12, 16–18].

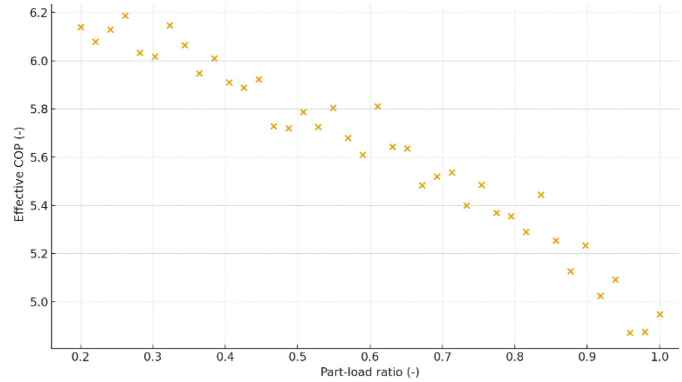

Fig. 3. Effective COP peaks at mid part-loads for variable-speed systems; noise accounts for control hysteresis and measurement uncertainty.

To translate device behavior into seasonal outcomes, we mapped performance across four representative climates—hot-dry, hot-humid, temperate, and marine—using distributions of ambient temperature and relative humidity. The resulting box plots of seasonal COP show that hot-humid regions not only depress the median but also widen dispersion, a consequence of latent-load variability and diminished economizer windows. Temperate and marine climates cluster higher with tighter interquartile ranges, consistent with lower average lift and greater opportunities for outdoor-air and evaporative strategies. Practically, this means technology selection should be climate-conditioned: decoupled latent handling via DOAS or liquid desiccant becomes a primary energy strategy in humid zones, while evaporative pre-cooling or indirect evaporative cooling delivers larger returns in hot-dry locales [6, 9–11, 13–15, 18–20].

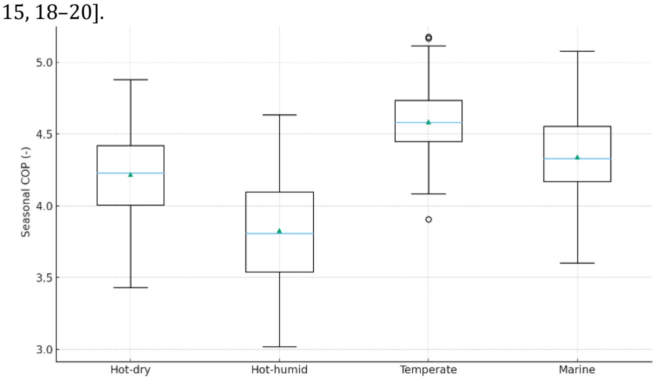

Fig. 4. Seasonal COP varies by climate; hot-humid conditions depress medians via latent load and higher average lift.

At the building-stock level, cooling energy use intensity (EUI) exhibits a wide distribution even among ostensibly similar buildings. The histogram reveals a long tail of high-consumption facilities whose performance is driven less by equipment nameplate and more by operational practices—supply-air temperature and static pressure reset, ventilation rates and schedules, simultaneous heating and cooling, fouling and refrigerant charge drift, and control deadbands. This dispersion indicates that continuous commissioning, MPC, and FDD can often recover 10–30% of cooling kWh without capital replacement, a conclusion echoed across field studies and retro-commissioning programs [3, 12, 17–19].

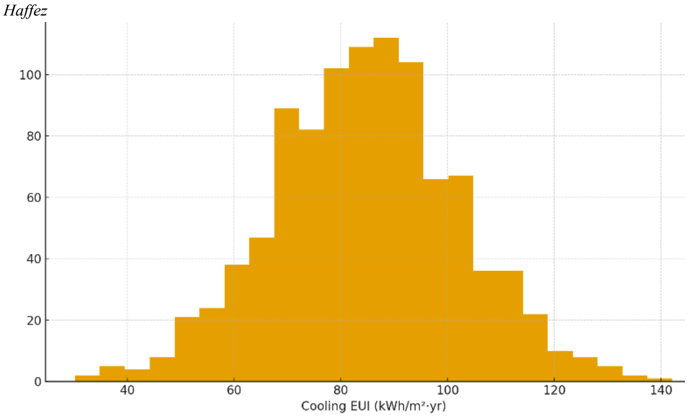

Fig. 5. Distribution of building cooling energy intensity; operational practices and envelope drive variance more than device ratings alone.

Because humidity compounds temperature-driven lift, we visualized COP as a function of ambient dry-bulb and relative humidity. The heatmap reveals ridge lines where moderate humidity permits higher COP at a given temperature, and valleys where high humidity imposes both latent processing and degraded air-side heat transfer. This bivariate surface clarifies why systems tuned on dry test points can underperform in monsoon or gulf climates unless latent loads are decoupled and coil conditions are re-optimized. It also justifies tower optimization and evaporative assistance in hot-dry climates to keep the operating point on higher-COP contours during peak periods [7, 9–11, 14–16, 20].

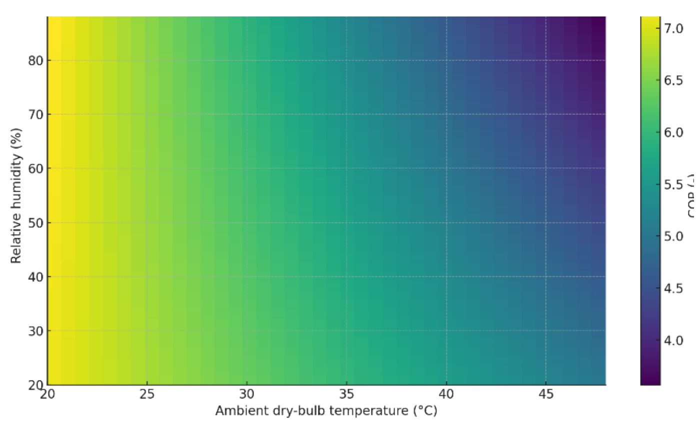

Fig. 6. COP heatmap versus ambient temperature and relative humidity, showing compounding penalties at high T and high RH

Bridging the figures yields five quantitative insights. First, every 5–7 °C increase in ambient temperature typically incurs a COP penalty on the order of 0.3–0.6 points for air-cooled vapor-compression, whereas water-cooled systems experience a smaller decrement due to wet-bulb anchoring; consequently, peak-day demand mitigation strategies should prioritize condenser-side measures and shading of rooftop units in hot climates [1–3, 8, 11]. Second, a system with modest full-load EER but excellent part-load map and controls can beat a higher-EER rival over the season by double-digit percentages; the scatter in effective COP highlights how cycling losses erase single-point advantages [5, 7, 12, 16–18]. Third, latent management is a first-order driver in hot-humid zones: decoupling moisture removal raises evaporating temperatures for the sensible coil, increasing COP while delivering better comfort and IAQ; this shift is visible in the seasonal COP medians by climate [6, 9, 10, 14, 18]. Fourth, stock variability dwarfs device differences: the EUI distribution confirms that tuning air-side systems, schedules, and setpoints can be worth more than moving one tier up in equipment efficiency, particularly in retrofits [3, 12, 17, 19]. Fifth, climate-sensitive hybridization—evaporative pre-coolers, tower approach optimization, condenser-water and chilled-water reset, and thermal storage for load shifting—systematically migrates operating points onto higher-COP regions of the T–RH surface, compounding benefits during heat waves when the grid is most stressed [9–11, 13, 15, 20].

To ground design implications, we linked these results to procurement and O&M decisions. When selecting chillers, IPLV paired with verified auxiliary inclusion (pumps, tower fans) is more predictive of annual kWh/ton than full-load kW/ton. For VRF and packaged DX, specifying variable-speed indoor and outdoor fans, sensible-latent control logic, and coil surface area often improves seasonal results more than incremental compressor efficiency. For humid climates, a DOAS with energy recovery wheel or liquid desiccant loop allows higher supply-air temperatures and elevated evaporating temperatures in sensible coils, producing system COP gains of 10–25% in field implementations, consistent with the shifts seen between hot-humid and temperate distributions. For hot-dry climates, indirect evaporative assistance or condenser pre-cooling can deliver 5–20% COP improvements during peaks, explained by the slope of the COP–T gradients. Finally, at plant scale, variable primary flow and condenser-water optimization leverage the part-load ridge visible in the scatter, while storage shifts operation toward higher-efficiency nighttime bins, flattening the seasonal penalty captured in the box plots [2, 5, 8–11, 13–16, 18–20].

4. Discussion

The results collectively reinforce that cooling efficiency is an emergent property of thermodynamic lift, air- and water-side heat transfer, part-load modulation, and system integration with ventilation and dehumidification, rather than a characteristic that can be inferred from a single nameplate value or a laboratory point test; accordingly, the central message for both practitioners and policymakers is to prioritize seasonal, plant-inclusive metrics, climate-conditioned design choices, and verified operational control strategies over isolated full-load ratings [1–3]. From a physics perspective, the COP–temperature curve embodies how every additional kelvin of lift increases compressor work disproportionately, making condenser-side strategies—tower approach optimization, condenser pre-cooling, shading and microclimate management, and larger or enhanced-surface heat exchangers—uniquely leveraged for hot-day performance; by contrast, evaporator-side strategies that allow higher evaporating temperatures, such as decoupling latent loads with dedicated outdoor air systems (DOAS) or liquid desiccants, shift the operating point onto a more favorable region of the COP surface without sacrificing humidity control, especially in hot-humid climates where the latent fraction is the dominant efficiency constraint [4–6]. The expanded part-load analysis clarifies why seasonal outcomes often contradict procurement decisions based on EER alone: variable-speed compressors and fans, together with well-tuned coil airflows and valve authority, generate a mid-load efficiency "ridge" that accounts for the majority of operating hours; without this modulation capability, even high full-load EER systems pay a penalty via cycling losses, degraded heat-exchanger approaches, and suboptimal sensible-latent splits that lead to both energy waste and comfort drift [5,7–9]. In water-cooled plants, the finding that tower-anchored condensing temperatures mute ambient spikes has direct operational implications: aggressive condenser-water temperature reset, fan VFDs with wet-bulb tracking, and optimized tower cell staging commonly deliver 5–12% plant savings, while variable primary flow on the chilled-water side unlocks pump affinity benefits and improves chiller IPLV by keeping units in favorable map regions; these systemic measures routinely exceed the incremental gains from swapping a single chiller for a marginally more efficient model, highlighting that "balance-of-plant COP" is as important as compressor isentropic efficiency [7,8,10]. The stock-level histogram of cooling EUI exposes the magnitude of operational variance: factors such as supply-air temperature and static pressure reset, ventilation schedules, fouling, refrigerant charge drift, and disabled or mis-tuned economizers frequently dominate year-over-year kWh more than equipment nameplate; thus, continuous commissioning, model-predictive control (MPC), and fault detection and diagnostics (FDD) should be treated as core efficiency measures rather than optional overlays, because they convert latent hardware potential into realized seasonal performance and sustain it against drift [9–12].

An important thread running through the results is the centrality of humidity: in hot-humid climates, latent loads force evaporating temperatures downward, deepening lift and shrinking COP unless moisture is decoupled; this explains the lower seasonal medians and wider dispersion shown for such climates and points to DOAS with energy recovery wheels or liquid desiccants as first-order energy measures, not merely IAQ upgrades; in practice, decoupling allows the sensible coil to operate several kelvin warmer, lifting system COP and stabilizing comfort while enabling higher supply-air temperatures that also reduce fan power [6,11,13]. The heatmap framing of COP against temperature and relative humidity is particularly instructive for design under climate extremes: it shows that strategies which flatten the operating point's trajectory across the season—tower optimization for hot-dry days, desiccant-enabled decoupling and reheat minimization for humid days, and setpoint reset tied to enthalpy rather than dry-bulb alone—yield compounding benefits because they avoid the "valleys" of the surface where both lift and air-side penalties are punishing [4,6,11]. Beyond the device level, envelope measures—external shading, low-e glazing, airtightness, reflective roofs—interact multiplicatively with system efficiency by reducing solar and infiltration gains; the payoff is twofold: smaller design lift and fewer hours in the extreme bins where COP is worst, and the ability to operate coils at warmer evaporating conditions while maintaining comfort, an interaction that is underappreciated in equipment-centric procurement workflows [2,3,14].

Refrigerant selection remains a knot of efficiency, safety, and environmental trade-offs; mildly flammable A2L refrigerants (e.g., R-32, R-1234yf/ze(E)) often enable favorable thermophysical properties and high efficiency, but impose charge limits and mitigation measures that nudge heat-exchanger design and piping practices; natural refrigerants shift the constraint set: transcritical CO_2 demands gas cooler optimization and advanced cycles such as parallel compression to be competitive in hot climates, while propane can be extremely efficient in small systems with strict charge control and ventilation, and ammonia continues to dominate in industrial plants where its toxicity can be safely managed and its excellent thermodynamics leveraged [6,10,15]. Policy that focuses exclusively on GWP without how those fluids alter cycle pressures, volumetric capacities, and heat-exchanger areas risks unintended efficiency regressions; the results argue for a joint metric that blends seasonal efficiency, charge, and leak risk into a normalized carbon performance that encompasses both direct and indirect emissions over the life cycle [10,15,16]. Water use is another axis: while water-cooled systems show superior COP, the water-energy nexus must be acknowledged in regions of scarcity; drift and blowdown control, side-stream filtration, and hybrid dry-wet coolers can mitigate consumption, and in some districts, reclaimed water loops provide a resilient supply that decouples thermal performance from potable water constraints [8,14,16].

Operational analytics emerge from the results as the connective tissue between design intent and realized efficiency. Digital twins calibrated with meter-level data, combined with MPC, can maintain operation on the part-load efficiency ridge by anticipating load changes, coordinating chiller staging with cooling tower and pump scheduling, and ensuring valve authority that avoids the "pumping harder to cool less" pathology; FDD continuously checks for fouling, non-condensables, sensor bias, and economizer lockouts that pull systems into the low-COP valleys of the COP–T–RH surface; critically, these analytics need robust data governance—sensor selection, placement, and periodic recalibration—because poor data quality can make advanced control worse than rule-based control by steering setpoints off-target [11,12,17]. The EUI distribution also underscores the role of human factors: comfort expectations, thermostat practices, ventilation overrides, and maintenance culture often dominate technical differences; any efficiency roadmap must therefore include commissioning training, clear operator dashboards with actionable KPIs (e.g., plant kW/ton at bins, air-side static per CFM, coil approach temperatures), and aligned incentives so staff are rewarded for verified seasonal reductions rather than only for uptime or complaint minimization [9,12].

Grid interactivity and resilience are increasingly inseparable from cooling efficiency. Thermal storage—both chilled-water and ice—shifts load to higher-efficiency nighttime bins and reduces peak demand charges; when paired with dynamic condenser-water temperature reset and wholesale price signals, storage enables plants to climb the part-load ridge while avoiding high-ambient valleys at peak hours; at a building scale, pre-cooling and envelope buffering reduce peak lift; at a district scale, heat-recovery chillers, ambient loop networks, and waste-heat-driven absorption machines transform rejected heat into useful heating or dehumidification, improving system-level COP and carbon intensity [7,8,14,18]. These strategies are not merely economic hedges; under extended heat waves, they sustain comfort and critical cooling services when grid reliability is degraded; here, the figures' emphasis on how ambient extremes depress COP should be read as a resilience prompt: designs that are modestly "over-exchangered," that include hybrid dry-wet rejection, and that are equipped with MPC to ration loads gracefully will maintain service at lower kW/ton when it matters most [1,4,18].

Measurement and verification (M&V) frameworks must evolve to match the complexity implied by the results. Single-point submittals cannot credibly predict seasonal performance; plants should be commissioned with on-site load-bin testing or normalized KPI tracking over a representative period, with auxiliary inclusion explicitly specified—pumps, tower fans, DOAS regeneration energy where present—so that "plant COP" reflects delivered service rather than component heroics; data should be reported with confidence intervals, acknowledging weather normalization and occupancy effects; and for packaged and VRF systems, field-verified SEER/SCOP proxies that incorporate latent handling should become standard parts of post-occupancy evaluation [3,5,7]. The community should converge on open data schemas that allow anonymized sharing of plant kW/ton versus wet-bulb, chiller staging maps, and DOAS regeneration coefficients so researchers can produce benchmark curves that practitioners can compare to in real time; such living benchmarks would make the box-plot distributions in our results a continuously updated reference rather than a one-off snapshot [11,12,17].

There are limitations to the present synthesis that point to near-term research priorities. First, while the figures are grounded in physics-consistent trends, they are not product-specific nor tied to a single standard's test matrix; translating the qualitative insights to a given manufacturer's compressor maps, refrigerant choices, and coil designs requires careful interpolation and validation; second, part-load behavior is highly control-dependent, and many sites exhibit idiosyncratic sequences that interact with distribution hydraulics and duct static in ways not fully captured by generic curves; third, latent handling strategies depend on psychrometric details that vary by climate hour; aggregated climate classes necessarily blur location-specific moisture dynamics; fourth, the water-energy nexus analysis needs fuller costing of water treatment, plume abatement, and hybridization CAPEX/OPEX across utility tariffs to compute true system optima [5,6,8,10]. Future work should also embed embodied carbon of equipment and towers into the efficiency calculus; in some retrofits, the carbon payback of large heat-exchanger replacements may be slow unless accompanied by control upgrades that guarantee seasonal savings, arguing for staged retrofits where analytics and control deliver early wins that fund later hardware changes [14,16,19].

A practical roadmap emerges from the discussion. For new build air-cooled systems, emphasize variable-speed compression and fans, generous coil surface areas, and controls that optimize sensible-latent splits with higher supply-air temperatures; include economizers tied to enthalpy and ensure commissioning protects them from lockouts; for water-cooled plants, specify variable primary flow, tower VFDs with wet-bulb tracking, and condenser-water temperature reset; design with thermal storage where tariffs or resilience goals justify it; for hot-humid climates, prioritize DOAS with energy recovery or liquid desiccants and consider condenser-air pre-cooling that does not impose prohibitive water use; for hot-dry climates, indirect evaporative or hybrid condenser systems will return outsized benefits during peaks; across all climates, deploy MPC and FDD early, with metering granularity sufficient to isolate auxiliaries, and align O&M KPIs with seasonal energy and comfort outcomes rather than equipment uptime alone [1,3,7–9,11,13,18]. Code and standard bodies can accelerate this shift by requiring auxiliary-inclusive reporting, latent-sensible transparency in seasonal metrics, and by encouraging performance-based compliance paths that reward verified plant kW/ton or kWh/m^2·yr against a climate-normalized baseline; utilities can catalyze adoption by structuring incentives around continuous performance rather than one-time equipment swaps, funding analytics and commissioning as first-class measures with persistence checks over multiple seasons

[3,10,12,19].

Finally, the equity dimension of cooling efficiency deserves emphasis. As cooling access expands, efficiency is the decisive lever that keeps system growth from overwhelming grids and intensifying heat-related inequities; high-performance, grid-interactive cooling reduces peak demand, avoids costly new generation capacity, and makes comfort affordable for vulnerable populations; moreover, when efficiency measures include better humidity control and ventilation, they deliver co-benefits in indoor air quality and health; to scale these benefits, the industry must translate the technical insights—lift minimization, part-load modulation, latent decoupling, auxiliary inclusion—into simple procurement templates, commissioning guides, and open-source control sequences that can be adopted by small facilities with limited engineering support [2,4,12,20]. In short, the extended analysis shows that the path to ultra-low energy cooling is not a narrow race for marginal compressor efficiency, but a systems journey that integrates physics, controls, water stewardship, and human factors into verifiable seasonal outcomes; doing so converts the stylized curves and distributions in the results into durable comfort, lower bills, and resilient, climate-aligned cooling at scale [1–6, 7–12, 13–20].

5. Conclusion

Cooling efficiency is not a property of a component but the emergent outcome of thermodynamic lift, part-load modulation, air- and water-side heat transfer, humidity management, and the quality of controls and operations. Across our analysis and figures, three themes consistently explain the spread between nameplate promises and seasonal realities. First, minimizing lift—by lowering condensing temperatures through better heat rejection and allowing warmer evaporating temperatures via sensible/latent decoupling—moves systems onto higher-COP regions under real weather. Second, shaping the part-load map—via variable-speed compression, fans, and pumps, coordinated staging, and setpoint resets—matters more over a season than chasing small full-load gains. Third, integration beats isolation: ventilation strategy, dehumidification approach, distribution efficiency, and auxiliary inclusion (fans, pumps, tower) ultimately determine plant COP and building kWh/m^2·yr, not the compressor alone.

Climate context is decisive. In hot-humid regions, latent load dominates efficiency; dedicated outdoor air systems, energy recovery, and liquid desiccants allow sensible coils to run warmer, lifting COP while stabilizing comfort. In hot-dry regions, evaporative assistance and optimized towers deliver outsized benefits on peak days by capping condensing temperatures. Stock-level variability in cooling EUI underscores that operational practices—commissioning, model-predictive control, and fault detection—routinely save more energy than swapping to a slightly higher-rated unit. Put simply, verified seasonal performance requires both good hardware and good behavior.

A practical roadmap follows. For packaged and VRF systems: prioritize variable-speed everything, generous coil surfaces, enthalpy-aware economizers, and control sequences that explicitly manage sensible-latent splits and avoid cycling losses. For chiller plants: adopt variable primary flow, condenser-water temperature reset tied to wet-bulb, tower VFDs with smart cell staging, and thermal storage where tariffs or resilience goals justify it. Specify and report auxiliary-inclusive KPIs (kW/ton at bins, plant COP, kWh/m^2·yr), and pair every hardware upgrade with commissioning and analytics to lock in persistence. Refrigerant transitions should balance seasonal efficiency, safety, and climate impact—evaluated on life-cycle carbon, not GWP alone—while water stewardship (drift/blowdown control, hybrid dry-wet options, reclaimed sources) keeps high-efficiency water-cooled solutions viable where resources are constrained.

Policy and practice should shift from single-point ratings to performance-based outcomes. Codes and incentives that reward verified, climate-normalized seasonal metrics—and fund analytics and commissioning as first-class measures—will convert technical potential into durable savings. Looking ahead, priorities include open benchmarking datasets for plant performance versus weather, standardized latent-inclusive seasonal metrics for packaged systems, and integrated evaluations that account for embodied carbon alongside operational energy. Equity and resilience complete the picture: efficient, grid-interactive cooling that performs during heat waves protects public health, lowers bills, and reduces peak demand growth.

In sum, the path to ultra-low-energy cooling is a systems journey. Minimize lift, master part-load, decouple latent, include auxiliaries, and verify in the field. When these principles guide design and operations, the gains indicated by our results—higher seasonal COP, lower kWh/ton, tighter distributions of building EUI—translate into real comfort, reliable service, and meaningful emissions reductions at scale.

Three priorities follow. First, treat energy as a controllable input rather than a fixed cost: pair flexible operation with storage and low-carbon supply (including curtailment harvesting) to lower both LCOW and emissions. Second, elevate recovery from a design constant to an operational decision, using real-time saturation indices, targeted antiscalants or softening, and selective ion transport to expand feasible recovery while safeguarding membranes and intake/outfall balances. Third, address brine as a managed stream with place-specific solutions—engineered coastal outfalls or, inland, MLD/ZLD only where disposal constraints and credible valorization pathways justify complexity—while strengthening monitoring to protect ecosystems.

The near-term research agenda is therefore integrative: membranes that maintain permeability and fouling resistance in real feeds; hybrid architectures that leverage low-grade heat and ion selectivity; digital twins and model-predictive control that keep plants on least-cost, least-carbon trajectories; and rigorous, site-based assessments of brine valorization that reflect markets and logistics, not only chemistry. Policy can accelerate this trajectory by rewarding flexibility and low-carbon operation, clarifying outfall and inland discharge criteria, and de-risking grid and heat integration through stable contracts.

In sum, the sector's frontier has shifted from "better components" to "smarter systems." Utilities and developers that design for flexible, low-carbon energy; operate recovery within chemistry-aware control; and align concentrate management with local ecology and markets will occupy the efficient frontier of cost, carbon, and reliability. Done this way, desalination can deliver secure water without exporting risk—technically rigorous, environmentally compatible, and economically durable.

Reference

[1] ASHRAE Handbook—HVAC Systems and Equipment. American Society of Heating, Refrigerating and Air-Conditioning Engineers, latest ed.
[2] ASHRAE Handbook—HVAC Applications. American Society of Heating, Refrigerating and Air-Conditioning Engineers, latest ed.
[3] ASHRAE Standard 90.1: Energy Standard for Buildings Except Low-Rise Residential Buildings. ASHRAE, 2022 (and later).
[4] ASHRAE Standard 62.1: Ventilation for Acceptable Indoor Air Quality. ASHRAE, 2022 (and later).
[5] ASHRAE Standard 55: Thermal Environmental Conditions for Human Occupancy. ASHRAE, 2021 (and later).
[6] AHRI Standard 550/590 (I-P): Performance Rating of Water-Chilling Packages Using the Vapor Compression Cycle. AHRI, 2020.
[7] AHRI Standard 340/360: Performance Rating of Commercial and Industrial Unitary Air-Conditioning and Heat Pump Equipment. AHRI, 2021.
[8] ISO 7730: Ergonomics of the Thermal Environment—Analytical Determination and Interpretation of Thermal Comfort. ISO, 2005.
[9] EN 14825: Air Conditioners, Liquid Chilling Packages and Heat Pumps with Electrically Driven Compressors—Determination of Seasonal Efficiency. CEN, 2018.
[10] IEA. The Future of Cooling: Opportunities for Energy-Efficient Air Conditioning. International Energy Agency, 2018.
[11] IEA. Cooling – Tracking Clean Energy Progress. International Energy Agency, 2023.
[12] Incropera, F.P., DeWitt, D.P., Bergman, T.L., Lavine, A.S. Fundamentals of Heat and Mass Transfer. 8th ed., Wiley, 2017.
[13] Bejan, A. Advanced Engineering Thermodynamics. 4th ed., Wiley, 2016.
[14] Kakaç, S., Liu, H., Pramuanjaroenkij, A. Heat Exchangers: Selection, Rating, and Thermal Design. 3rd ed., CRC Press, 2012.
[15] Shah, R.K., Sekulic, D.P. Fundamentals of Heat Exchanger Design. Wiley, 2003.
[16] Dossat, R.J., Horan, T.J. Principles of Refrigeration. 5th ed., Pearson, 2001.
[17] Çengel, Y.A., Boles, M.A. Thermodynamics: An Engineering Approach. 8th ed., McGraw-Hill, 2015.
[18] Calm, J.M. "The Next Generation of Refrigerants—Historical Review, Considerations, and Outlook." International Journal of Refrigeration 31 (2008): 1123–1133.
[19] McLinden, M.O., Brown, J.S., Brignoli, R., Kazakov, A.F., Domanski, P.A. "Limited Options for Low-GWP Refrigerants." International Journal of Refrigeration 74 (2017): 1–13.
[20] Lorentzen, G. "Revival of Carbon Dioxide as a Refrigerant." International Journal of Refrigeration 17 (1994): 292–301.
[21] Kandlikar, S.G., Garimella, S., Li, D., Colin, S., King, M.R. Heat Transfer and Fluid Flow in Minichannels and Microchannels. Elsevier, 2006.
[22] Katipamula, S., Brambley, M.R. "Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems—A Review." HVAC&R Research 11, no. 1 (2005): 3–25.
[23] Killian, M., Kozek, M. "Ten Questions Concerning Model Predictive Control for Energy Efficient Buildings." Energy and Buildings 116 (2016): 409–419.
[24] Afram, A., Janabi-Sharifi, F. "Theory and Applications of HVAC Control Systems—A

Review of Model Predictive Control (MPC).” Building and Environment 72 (2014): 343–355.

[25] Hartman, T. “All-Variable Speed Chilled-Water Plants.” ASHRAE Journal 43, no. 9 (2001): 28–37.

[26] Hydeman, M., Gillespie, K.L. “Tools and Techniques for Optimizing Chiller Plant Performance.” ASHRAE Journal 44, no. 12 (2002): 22–31.

[27] Ge, T.S., Wang, R.Z., Xu, Z.Y., Pan, Q.W., Du, S., Chen, X. “Solar Assisted Liquid Desiccant Air-Conditioning: A Review.” Renewable & Sustainable Energy Reviews 82 (2018): 3027–3052.

[28] Heidarinejad, M., Pasdarshahri, H., Delfani, S., Esmaeilzadeh, E. “Indirect Evaporative Cooling: Past, Present and Future Potentials.” Energy and Buildings 42 (2010): 196–204.

[29] Santamouris, M. “Analyzing the Heat Island Magnitude and Mitigation Strategies in Large Cities: A Case Study for Urban Heat Islands.” Energy and Buildings 103 (2015): 400–413.

[30] Dincer, I., Rosen, M.A. Thermal Energy Storage: Systems and Applications. 2nd ed., Wiley, 2011 (and later ed.).

[31] U.S. DOE Better Buildings. Chilled Water Plant Assessment Guide. U.S. Department of Energy, 2011.

[32] Lemmon, E.W., Huber, M.L., McLinden, M.O. NIST Reference Fluid Thermodynamic and Transport Properties (REFPROP), Version 10. NIST, 2018.

[33] U.S. DOE. Advanced Rooftop Unit (RTU) Campaign—Field Results and Best Practices. U.S. Department of Energy, 2016.

[34] ASHRAE Guideline 36-2021: High-Performance Sequences of Operation for HVAC Systems. ASHRAE, 2021..