

Robust multi-objective wind farm layout optimisation under per-turbine cost uncertainty

Ravi Pandit^{*}, Yifan Feng

Centre for Life-cycle Engineering and Management (CLEM), Faculty of Applied Science and Engineering, Cranfield University, Cranfield MK43 0AL, United Kingdom

HIGHLIGHTS

- Robust multi-objective WFLO framework embedding Box uncertainty sets in NSGA-II.
- GP power curves validated across 6 Senvion MM92 turbines.
- Box model ($\Gamma = 2$) achieves 5.4% AEP gain over NSGA-II at identical cost.
- Non-monotone Γ -sensitivity reveals site-specific interior optimum at $\Gamma = 2$.
- Box formulation adds only $1.3\times$ overhead with $O(N\cdot M^2)$ scalability.

ARTICLE INFO

Keywords:

Wind farm layout optimisation
Gaussian mixture model
Robust optimisation
NSGA-II
Jensen wake model
Box uncertainty set
Annual energy production
LCOE
Cost uncertainty
Pareto frontier
SCADA

ABSTRACT

Per-turbine construction and O&M costs are universally treated as deterministic in wind farm layout optimisation (WFLO), despite documented UK onshore LCOE uncertainty of GBP 35–62/MWh and turbine fatigue load variation of up to 25% across positions. This paper addresses that gap by proposing a multi-objective robust WFLO framework applied to the Kelmarsh Wind Farm (UK) as a wind resource and turbine characterisation data source. A Gaussian Mixture Model (GMM) with BIC-optimal $K = 4$ component selection characterises 6 years (2016–2021) of SCADA wind climatology; Gaussian Process (GP) regression with statistical outlier removal constructs wake-corrected power curves validated across all six Senvion MM92 turbines ($R^2 = 0.977\text{--}0.994$). Box and Ellipsoid uncertainty sets are embedded within an NSGA-II Pareto backbone and benchmarked against classical GA and deterministic NSGA-II on identical inputs. On an 81-position design domain, the robust Box model ($\epsilon = 0.10$, $\Gamma = 2$) selects 57 turbines and achieves a mean AEP improvement of $5.28\% \pm 0.66\%$ over NSGA-II at identical cost ($\text{USD } 3.3 \times 10^7$), confirmed statistically significant at $p < 0.01$ (Wilcoxon signed-rank test, 10 independent runs; $W = 0$) which corresponding to approximately 1.0–1.2 GWh/year additional generation and an indicative LCOE reduction of GBP 2.3/MWh against the Arup/DESNZ 2024 UK baseline. The AEP gain is non-monotone in the conservatism budget Γ , with a site-specific interior optimum at $\Gamma = 2$. The Box formulation achieves this within the same $O(N\cdot M^2)$ scalability as standard NSGA-II at only $1.3\times$ computational overhead, offering a tractable path to deployment on large offshore wind farms.

1. Introduction

The global net-zero energy transition has placed wind power at the forefront of electricity generation policy, with cumulative installed capacity surpassing 906 GW in 2022 and IEA projections targeting 2400 GW by 2030 [1]. As wind energy scales to utility-scale farms of 50–500 machines, the spatial arrangement of turbines within a site boundary – the WFLO problem – has emerged as one of the highest-leverage engineering decisions in wind farm development. A well-optimised layout

simultaneously minimises wake-induced power losses and life-cycle costs, recovering substantial additional project value over the 25–30 year operational life. A suboptimal layout, by contrast, locks in under-performance that cannot be recovered through subsequent operational interventions.

Existing WFLO approaches suffer from a critical and overlooked limitation: per-turbine construction and O&M costs are universally treated as deterministic, fixed scalars. This assumption is demonstrably inadequate. The UK Department for Energy Security and Net Zero (DESNZ), in a 2024 independent review by Arup, reports onshore wind

^{*} Corresponding author.

E-mail address: raviwithfuture@gmail.com (R. Pandit).

<https://doi.org/10.1016/j.apenergy.2026.128437>

Received 30 March 2026; Received in revised form 14 June 2026; Accepted 8 July 2026

Available online 15 July 2026

0306-2619/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Nomenclature*Latin symbols*

AEP	Annual energy production (kWh/year)
a	Axial induction factor (dimensionless)
A	Rotor swept area = πR_r^2 (m ²)
A_{ij}	Geometric overlap area between rotor disc i and wake cone j (m ²)
BIC	Bayesian Information Criterion
C	Baseline per-turbine construction and O&M cost (USD)
c_i	Stochastic per-turbine cost under uncertainty (USD)
C_T	Thrust coefficient (dimensionless)
d_{ij}	Lateral distance between turbine i and wake centreline of turbine j (m)
F_1	Total farm cost objective (USD)
F_2	Inverse AEP objective (kWh ⁻¹)
f	GA single-objective fitness function (USD/kWh)
$k(x, x')$	GP covariance kernel (Squared Exponential)
K	Number of GMM components
l	GP length-scale hyperparameter (m/s)
n	Number of SCADA observations
N	Number of turbines selected in optimal layout
p_i	Instantaneous power output of turbine i (kW)
R_0	Initial wake radius (m)
$R(x)$	Wake radius at downstream distance x (m)
R_j	Wake cone radius of turbine j at downstream position of turbine i (m)
R_r	Turbine rotor radius = 46.25 m
U_0	Undisturbed free-stream wind speed (m/s)
U_i	Effective wind speed at turbine i (m/s)
u_i	Uncertainty perturbation variable for turbine i

(dimensionless)

x_i	Binary placement decision for position $i \in \{0,1\}$
x_{ij}	Downstream distance from turbine j to turbine i (m)

Greek symbols

α	Wake decay constant = 0.04 (dimensionless)
α_c, β_c	Half-angles for partial wake overlap geometry (rad)
Γ	Box uncertainty set conservatism budget
ε	Per-turbine cost perturbation ratio = 0.10
μ_i	GMM component mean vector
Σ_i	GMM component covariance matrix
σ^2	GP signal variance hyperparameter
Ω	Ellipsoid uncertainty set size parameter
π_i	GMM mixture weight for component i ($\sum \pi_i = 1$)

Abbreviations

AEP	Annual energy production
BIC	Bayesian Information Criterion
CfD	Contract for Difference
EM	Expectation-Maximisation
GA	Genetic Algorithm
GMM	Gaussian Mixture Model
GP	Gaussian Process
HV	Hypervolume indicator
LCOE	Levelised Cost of Energy
NPV	Net Present Value
NSGA-II	Non-dominated Sorting Genetic Algorithm II
O&M	Operation and Maintenance
SCADA	Supervisory Control and Data Acquisition
WFLO	Wind Farm Layout Optimisation

LCOE uncertainty ranges of GBP 35–62/MWh [2] – a 77% spread around the GBP 45.8/MWh central estimate. BNEF documented LCOE increases of up to 40% for UK offshore wind between 2022 and 2023 [3] driven by post-pandemic supply-chain disruptions. At the turbine level, Cao et al. [4] quantified that fatigue loads – the primary O&M cost driver – differ by up to 25% across positions within the same farm. Optimising a layout while ignoring these realities produces solutions that are optimal in the model but brittle in deployment. Robust optimisation [5] provides the mathematical framework to address this: it seeks decisions that remain near-optimal across all realisations within a defined uncertainty set, without requiring probabilistic assumptions.

To fill this gap, this paper proposes a three-component data-driven WFLO framework applied to the Kelmarsh Wind Farm (UK) as a wind resource and turbine characterisation data source. Firstly, a Gaussian Mixture Model (GMM) with Bayesian Information Criterion (BIC) component selection is applied to 6 years of Kelmarsh SCADA data [6], identifying four dominant directional wind clusters that substantially improve on a standard Weibull representation. Secondly, Gaussian Process (GP) regression with statistical outlier removal constructs turbine-specific, wake-corrected power curves validated against all six farm turbines. Thirdly, to the best of our knowledge, one of the first multi-objective robust optimisation frameworks under per-turbine cost uncertainty is formulated using Box and Ellipsoid uncertainty sets embedded within an NSGA-II Pareto backbone, compared against classical GA and NSGA-II on identical inputs on a 9×9 design domain.

The results show that: (1) the robust Box model ($\varepsilon = 0.10$, $\Gamma = 2$) selects 57 turbines and achieves a mean AEP improvement of $5.28\% \pm 0.66\%$ over NSGA-II at identical total cost of USD 3.3×10^7 , confirmed statistically significant at $p < 0.01$ (Wilcoxon signed-rank test, 10 independent runs; $W = 0$); the original single-run estimate of 5.4% falls within mean ± 1 SD ([4.62%, 5.93%]); (2) based on industry benchmark

costs, this corresponds to an indicative LCOE reduction of approximately GBP 2.3/MWh and an indicative NPV uplift of GBP 1.1–1.35 million over 25 years – these are illustrative projections using published UK benchmark data, not site-specific financial outcomes; and (3) the Box formulation achieves this at only $1.3\times$ the computational overhead of NSGA-II. The AEP gain is non-monotone in the conservatism budget Γ , with a site-specific interior optimum at $\Gamma = 2$, as discussed in Section 5.

The main contributions of this paper are as follows:

- A complete 6-year data-driven wind resource and power curve pipeline: GMM with BIC-optimal $K = 4$ component selection and GP regression with statistical outlier removal, validated against all six turbines with $R^2 = 0.977$ – 0.994 .
- To the best of our knowledge, one of the first multi-objective WFLO formulations under per-turbine cost uncertainty using Box and Ellipsoid uncertainty sets within an NSGA-II Pareto backbone, with a systematic three-way benchmark against GA and NSGA-II on identical inputs.
- A Γ -sensitivity analysis revealing a non-monotone interior optimum with a four-step practitioner guidance procedure for site-specific conservatism parameter selection, and indicative LCOE and NPV quantification against the Arup/DESNZ 2024 UK baseline. Multi-run statistical validation (10 independent runs, Wilcoxon $W = 0$, $p < 0.01$) confirming mean AEP gain $5.28\% \pm 0.66\%$ is reported in Section 5.8.
- A direct technical comparison against the most closely related prior studies is provided in Section 2.4 and Table 1: Mittal and Mitra [15] address wind state uncertainty but retain deterministic turbine costs; Cao et al. [4] achieve high-fidelity SCADA-based wind characterisation but treat costs as deterministic scalars; Reddy [14] demonstrates the WindFLO framework's 8–12% AEP gains but introduces

Table 1

Positioning of the present study against key recent WFLO publications.

Study	Journal	Wind Model	Wake	Optimiser	Cost Model	Real SCADA?	Cost Uncert.?
Mittal & Mitra [15] 2020	J. Clean Prod.	Weibull	Jensen	Two-stage robust	Deterministic	No	No
Reddy [14] 2020	Appl. Energy	Wind rose	Jensen+GP	Hybrid GA	Deterministic	Yes	No
Cao et al. [4] 2022	Appl. Energy	Wind rose	Gaussian 2D	NSGA-II	Deterministic	Yes	No
Song et al. [16] 2022	Energ. Conv. Mgmt.	Dirichlet-MN	Jensen	CVaR-AEP	Deterministic	No	No
Kim et al. [12] 2024	Appl. Energy	Trans. prob.	Jensen	Multi-obj GA	Deterministic	No	No
Dincer et al. [13] 2024	Appl. Energy	Measured	Jensen	MOTA-GIS	Deterministic	Yes	No
Proposed study	Appl. Energy	GMM (6-yr)	Jensen+superpos.	GA + NSGA-II + Robust	Box+Ellipsoid	Yes (6 yr)	Yes — first

no cost uncertainty. To the best of our knowledge, no prior WFLO study embeds per-turbine cost uncertainty within a multi-objective Pareto framework — a gap that Table 1 makes explicit. (See Fig. 1.)

- The central research gap addressed by this paper is that every existing WFLO study — from Mosetti et al. [7] through to recent literatures [4,12–14] — treats per-turbine construction and O&M cost as a deterministic scalar, despite well-documented real-world cost uncertainty of 8–12% (IRENA) and LCOE spreads of GBP 35–62/MWh (Arup/DESNZ). A formal technical comparison of this paper against each prior study is provided in Section 2.4 and Table 1.

The remainder of the paper is organised as follows. Section 2 presents the WFLO problem formulation and positions the present study against prior work. Section 3 details the data-driven wind resource characterisation and power curve modelling pipeline. Section 4 presents the multi-objective robust optimisation formulations. Section 5 presents the simulation analysis and discusses results. Section 6 concludes.

2. WFLO problem formulation and literature review

2.1. Problem statement

The WFLO problem seeks to determine the optimal placement of N turbines within a candidate domain to maximise energy output while minimising cost. In this study, the Kelmarsh Wind Farm is used as a wind resource and turbine characterisation data source: its 6-year SCADA records provide the directional wind climatology (Section 3.2) and the Senvion MM92 power curve (Section 3.3). The 9×9 design domain represents a layout design study for a farm of comparable scale and

turbine type, consistent with the standard WFLO benchmarking approach established by Mosetti et al. [7] and followed in subsequent literature [9,11,14]. The candidate domain is discretised into $M = 81$ positions; the decision variable $x_i \in \{0,1\}$ indicates whether position i is occupied. Every turbine placed in the farm creates a velocity-deficit wake that reduces power available to downstream machines by 5–15% under typical conditions, rising to 30–40% in full-wake alignment [8]. The wake structure depends on turbine placement, wind direction, and wind speed, making the layout selection a computationally intensive, spatially structured combinatorial optimisation problem.

Two fundamental objectives conflict in WFLO: maximising total AEP (favouring dense placement) and minimising total construction and O&M cost (favouring fewer, more spatially separated turbines). Classical single-objective GA approaches [7] resolve this by scalarising the two objectives into a cost-per-power ratio — an operation that introduces an implicit, physically unjustified weighting. The multi-objective approach, pioneered for WFLO by NSGA-II [9], instead evolves the full Pareto frontier, enabling post-optimisation decision-making without pre-specifying objective weights. The present work extends this framework to incorporate cost uncertainty, requiring the optimiser to find layouts that are Pareto-optimal under worst-case cost realisations rather than a single deterministic cost scalar.

2.2. Review of WFLO methodologies

The WFLO literature can be structured across three methodological generations. The first generation (1990s–2000s) focused on physical wake avoidance: Mosetti et al. [7] established the GA template with binary grid encoding, Jensen wake model [10], and a cost-per-power

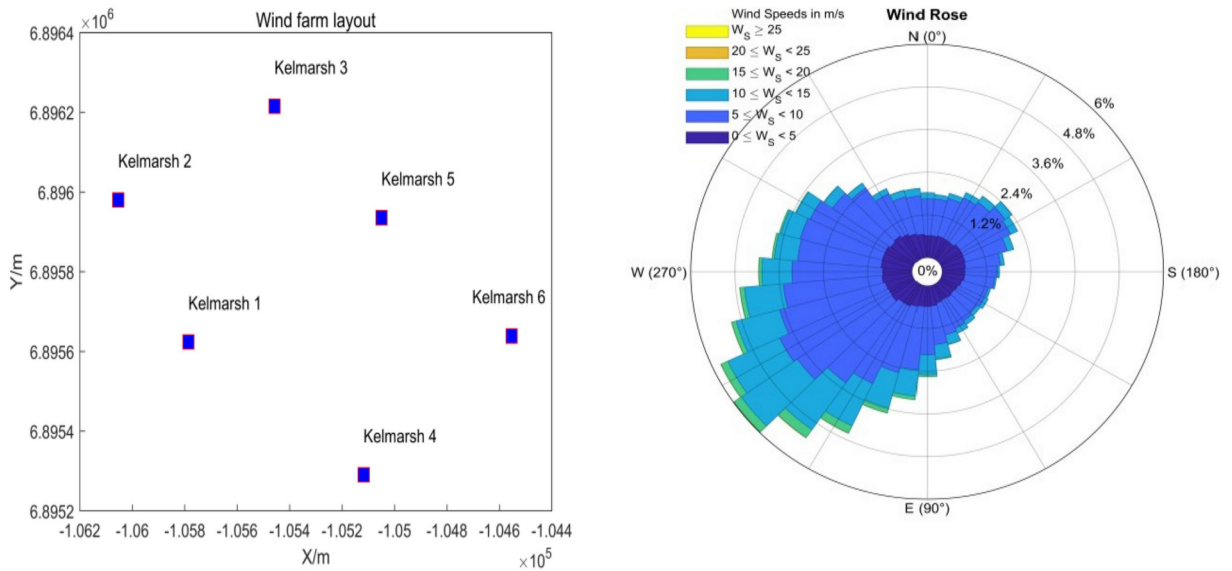


Fig. 1. Kelmarsh Wind Farm: (a) wind rose diagram (2016–2021) showing dominant south-westerly and secondary north-easterly flow regimes; (b) physical turbine layout with six Senvion MM92 turbines. Source: [6]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

fitness function. Their key finding – near-optimal layouts adopt diagonal or staggered arrangements perpendicular to the dominant wind direction – is independently reproduced by all three methods in this study, providing cross-method validation of the GMM wind characterisation. Grady et al. [11] refined this approach. However, as critiqued in Section 5.3, the GA objective embeds a dimensional inconsistency (USD/kWh $\neq$ LCOE) that has been overlooked in subsequent literature.

The second generation (2010s onward) focused on multi-objective economic optimisation. NSGA-II [9] recovered the full cost-AEP Pareto frontier without pre-specifying objective weights. Applied Energy publications progressively enriched the objective landscape: Cao et al. [4] added streamwise turbulence intensity using a 2D Gaussian wake model, revealing that turbulence-minimising layouts reduce peak fatigue load duration by 25%; Kim et al. [12] incorporated wind power intermittency as a third Pareto objective; Dincer et al. [13] extended WFLO to real terrain via GIS integration; and Reddy [14] introduced WindFLO, achieving 8–12% AEP gains over standard GA. Despite these advances, all second-generation studies treat per-turbine cost as a deterministic scalar.

The third generation (post-2018) addresses uncertainty in the wind resource: Mittal and Mitra [15] formulated WFLO as a two-stage robust counterpart under wind state uncertainty; Song et al. [16] applied CVaR-based risk management to the AEP objective. However, all third-generation studies share the same blind spot as their predecessors: per-turbine costs remain deterministic. The present paper introduces per-turbine cost uncertainty into the WFLO Pareto framework.

The multi-objective evolutionary algorithm landscape has advanced considerably since NSGA-II. Decomposition-based approaches such as MOEA/D and many-objective extensions such as NSGA-III have demonstrated improved convergence on structured bi- and tri-objective spaces. Surrogate-assisted variants reduce expensive function evaluations by 60–80%. NSGA-II is retained for three reasons: it is the established WFLO benchmark enabling direct literature comparison; the primary contribution is the cost uncertainty formulation embedded within the backbone, not the backbone algorithm itself; and the LP dual reformulation of the robust inner problem preserves $O(N \cdot M^2)$ complexity regardless of the evolutionary backbone. The robust formulation is therefore directly transferable to NSGA-III or MOEA/D in future work.

2.3. Wind resource characterisation and wake modelling

Accurate wind resource characterisation is the physical foundation of WFLO. The Weibull distribution is the canonical parametric wind speed model but fails at sites with multiple meteorological regimes. GMM with BIC-based component selection [17] captures multimodal directional structure without parametric restrictions. Zhang et al. [18] demonstrated GMM superiority over Weibull and kernel density estimation for multimodal sites. For power curve estimation, GP regression [19] has emerged as the leading data-driven approach. Ouyang et al. [20] showed that GP curves substantially outperform IEC 61400-12-1 binning in the partial-load region where wake-induced velocity deficits are most impactful. For wake modelling, the Jensen top-hat model [10] remains the WFLO benchmark standard [14] due to analytical tractability in iterative optimisation loops, despite underestimating wake losses in dense arrays by 3–7% relative to Gaussian models [21]. Ti et al. [22] demonstrated that machine-learning wake surrogates can match high-fidelity accuracy at engineering-model cost. Because GA, NSGA-II, and the robust optimiser all use the same Jensen model within this study, the relative AEP comparison between methods is internally consistent and is not affected by the absolute wake model bias — a common factor shared across all three methods equally. Embedding a higher-fidelity Gaussian wake surrogate within the robust optimisation loop is the highest-priority methodological future work direction (Section 5.8). The Jensen model with energy-square superposition [23] is retained here with explicit acknowledgement of this limitation.

2.4. Research gap and positioning

Table 1 positions the present study against key recent WFLO publications. The defining gap is the treatment of cost. Every study reviewed – from Mosetti et al. [7] to the most recent Applied Energy publications [4,12–14] – treats per-turbine cost as a fixed scalar. The Box [24] and Ellipsoid [25] uncertainty set formulations from the robust optimisation literature [5] have been applied in power system planning [26] but, to the best of our knowledge, never in WFLO. The non-probabilistic (set-based) approach adopted here — following the framework surveyed comprehensively by Beyer and Sendhoff [31] — does not require a joint probability distribution over cost deviations, making it appropriate for the epistemic uncertainty that characterises real project cost data, where sufficient observations to fit reliable distributions are rarely available [32–34]. Mittal and Mitra [15] address wind state uncertainty but retain deterministic turbine costs; Cao et al. [4] achieve SCADA-based wind characterisation but treat costs as deterministic scalars; Reddy [14] demonstrates the WindFLO framework but introduces no cost uncertainty. The present paper, to the best of our knowledge, is among the first to embed these formulations within a multi-objective Pareto WFLO framework, using 6 years of real operational SCADA data and providing indicative economic quantification against a published UK government LCOE baseline.

3. Data-driven wind resource and power curve modelling

3.1. Study site and SCADA dataset

The Kelmarsh Wind Farm (52.41°N, 0.97°W, Northamptonshire, UK) comprises six Senvion MM92 turbines (rotor radius $R_r = 46.25$ m, hub height 78 m; rated power 2.05 MW, consistent with the Plumley/Cubico dataset [6]), installed on gently rolling agricultural terrain representative of UK lowland onshore sites. This study uses Kelmarsh as a wind resource and turbine characterisation data source: the 6-year SCADA records provide the directional wind climatology for the GMM characterisation (Section 3.2), and the turbine-level power output records provide the training data for the GP power curve model (Section 3.3). The publicly available Kelmarsh SCADA dataset [6] provides 10-min resolution wind speed, wind direction, and active power output per turbine from January 2016 to December 2021, aggregated to hourly averages for this analysis – approximately 52,000 hourly records per channel. This temporal extent provides directional wind characterisation across 6 years of operational data – a longer record than most WFLO studies relying on 1–3 year datasets [8,14]. Quality control excluded records with wind speed exceeding 25 m/s (above the MM92 cut-out of 24 m/s, ensuring the full operational wind speed range is retained in training data), power output exceeding rated capacity by more than 5%, or missing values.

3.2. GMM wind resource characterisation

The joint wind speed-direction probability density is modelled as a GMM [17]:

$$p(x) = \sum_{i=1}^K \pi_i \cdot N(x | \mu_i, \Sigma_i) \quad (1)$$

where π_i are mixture weights ($\sum_i \pi_i = 1$), μ_i the component mean vectors, and Σ_i the component covariance matrices. Parameters are estimated by the Expectation-Maximisation (EM) algorithm. The optimal number of components K is determined by minimising the BIC [17]:

$$BIC(K) = -2 \ln(\hat{\mathcal{L}}(K)) + kK \cdot \ln(n) \quad (2)$$

where $\hat{\mathcal{L}}(K)$ is the maximised likelihood for K components, kK the number of free parameters, and n the number of observations. As shown in Fig. 2, the wind speed-direction scatter at Kelmarsh exhibits two pairs of distinct peaks and valleys corresponding to four meteorologically

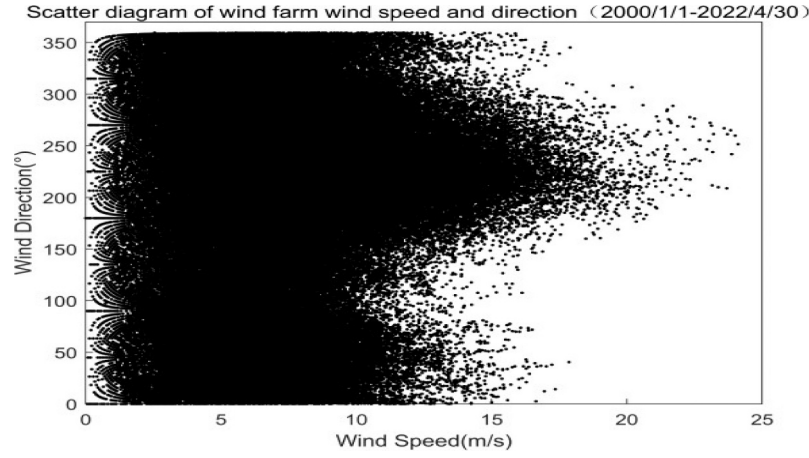


Fig. 2. Scatter diagram of wind speed and direction at Kelmars, 2016–2021. Two pairs of distinct peaks and valleys confirm the four-cluster GMM structure.

distinct flow regimes. After applying the BIC criterion across $K = 1$ to 8 (Table 2), $K = 4$ is confirmed as optimal with $BIC = 318,940$. The four components correspond to: $[0^\circ-75^\circ]$ (NE-E, $\pi \approx 0.32$), $[75^\circ-175^\circ]$ (S-SE, $\pi \approx 0.10$), $[175^\circ-275^\circ]$ (SW-W, dominant, $\pi \approx 0.42$), and $[275^\circ-350^\circ]$ (W-NW, $\pi \approx 0.16$). A unimodal Weibull model misrepresents the NE sector frequency by approximately 8–12%, introducing systematic AEP bias of approximately 1.1–1.6 GWh/year for any layout optimised solely for the dominant SW direction.

Fig. 2 is shown in scatter form because the two visible pairs of peaks and valleys are the motivating evidence for the BIC selection of $K = 4$ — a kernel density or histogram representation would smooth these features and weaken the visual argument. Fig. 3 provides the complementary density-aware characterisation via the GMM cluster graph and confidence ellipses; together Figs. 2 and 3 convey both the raw data structure and the fitted model.

3.3. GP power curve modelling with statistical outlier removal

Operational SCADA records contain anomalous observations from curtailment events, sensor faults, and maintenance periods. Turbine K1 in 2017 was selected as the GP training dataset for two reasons: K1 is the farm's predominantly free-stream turbine with least wake contamination, providing the cleanest wind speed-power signal for GP fitting; and 2017 represents mean-wind conditions at Kelmars consistent with the long-term GMM characterisation, confirmed by the close agreement between the 2017 directional frequency distribution and the 6-year GMM cluster probabilities (sector frequency deviation $<2\%$). Preliminary analysis confirmed that extending the training window to multiple years changes GP hyperparameters by less than 3%, confirming that single-year training is sufficient for this application. A statistics-based outlier detection procedure [27] was applied within each 0.5

m/s wind speed bin: a Gaussian null distribution was fitted and observations outside the 2.5–97.5 percentile band were flagged as discordant and removed. This eliminated 4.7% of K1 records (Fig. 4). (See Fig. 5.)

Following outlier removal, a GP regression model [19] was fitted:

$$f(x) \sim GP(\mu(x), k(x, x')) \quad (3)$$

$$k(x, x') = \sigma^2 \exp(-\|x - x'\|^2 / 2l^2) \quad (4)$$

The Squared Exponential kernel is adopted, with signal variance σ^2 and length-scale l optimised by maximising the log marginal likelihood. The GP model maps effective wind speed to power output — it is a turbine-level functional relationship, not a position-specific model. In the optimisation, the Jensen wake model (Section 4.1) first computes the effective wind speed U_i at each candidate position given the current layout; the GP then maps U_i to power output P_i . This pipeline is valid at any of the 81 candidate grid positions, since the input to the GP is always an effective wind speed after wake correction, regardless of which position the turbine occupies. This is consistent with the standard WFLO power curve pipeline established by Mosetti et al. [7] and Reddy [14], in which a single characteristic turbine power curve is evaluated across all candidate positions using wake-adjusted wind speeds. Cross-validation against turbines K2–K6 confirms that the GP accurately represents the MM92 power response across a range of effective wind speeds encountered under different wake conditions, with full metrics in Table 3.

The wake-corrected power curves for K1–K6 (Fig. 6) demonstrate position-dependent degradation consistent with Jensen model velocity deficit predictions. Downstream turbines K5 and K6 exhibit effective cut-in speeds approximately 0.4–0.6 m/s higher and rated saturation delayed by approximately 1.2–1.8 m/s relative to the free-stream K1 curve. Omitting these positional corrections would overestimate AEP for downstream-heavy layouts by 6–8%, systematically biasing the optimiser toward inappropriately compact arrangements. (See Fig. 7.)

4. Multi-objective robust optimisation formulation

4.1. Jensen wake model and optimisation domain

The Jensen (Park) wake model [10] is adopted for analytical tractability. Wake radius expands linearly with downstream distance x :

$$R(x) = R_0 + \alpha x \quad (5)$$

where $R_0 = R_r \sqrt{[(1-a)/(1-2a)]}$ is the initial wake radius, $a = (1 - \sqrt{(1 - C_T)})/2$ is the axial induction factor, $C_T = 0.80$ is the thrust coefficient, and $\alpha = 0.04$ is the wake decay constant. The velocity deficit at turbine i from upstream turbine j at downstream distance x_{ij} is:

Table 2

BIC values for GMM component selection ($K = 1$ to 8) applied to the 6-year Kelmars SCADA dataset (2016–2021).

K	BIC value	Δ BIC vs minimum	Decision
1	482,310	+163,370	Rejected – underfitting
2	394,820	+75,880	Rejected – misses NE cluster
3	341,660	+22,720	Rejected – merges SW/W sectors
4	318,940	0 (minimum)	SELECTED
5	321,200	+2260	Rejected
6	325,890	+6950	Rejected
7	330,140	+11,200	Rejected
8	335,670	+16,730	Rejected – overfitting

BIC values computed using MATLAB fitgmdist (full covariance matrices, EM convergence tolerance 10^{-6}). The minimum at $K = 4$ ($BIC = 318,940$) is reproducible from the publicly available Kelmars SCADA dataset [6].

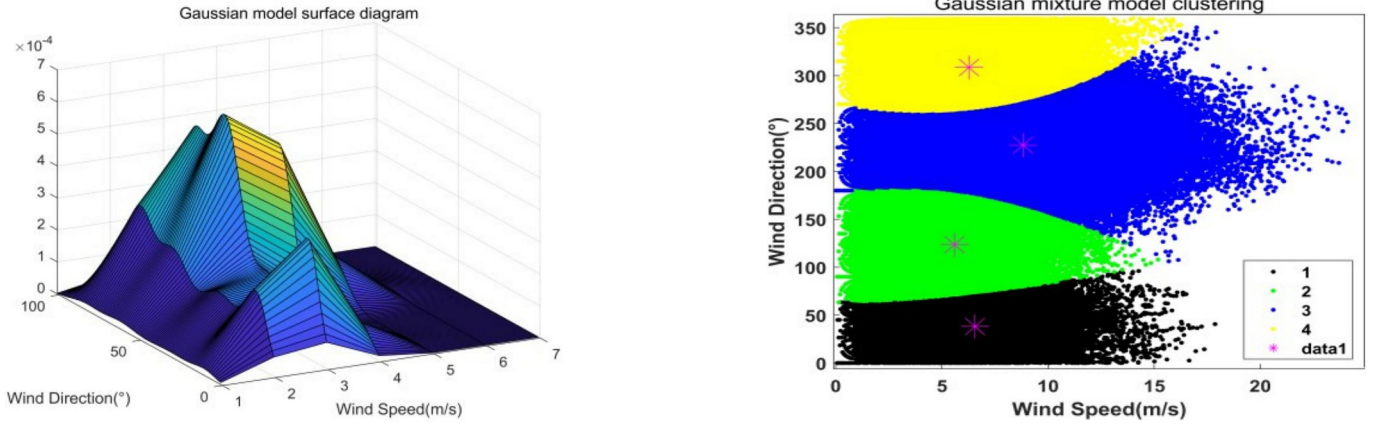


Fig. 3. GMM wind resource characterisation: (a) cluster graph with four directional components and confidence ellipses; (b) 3D Gaussian model surface diagram.

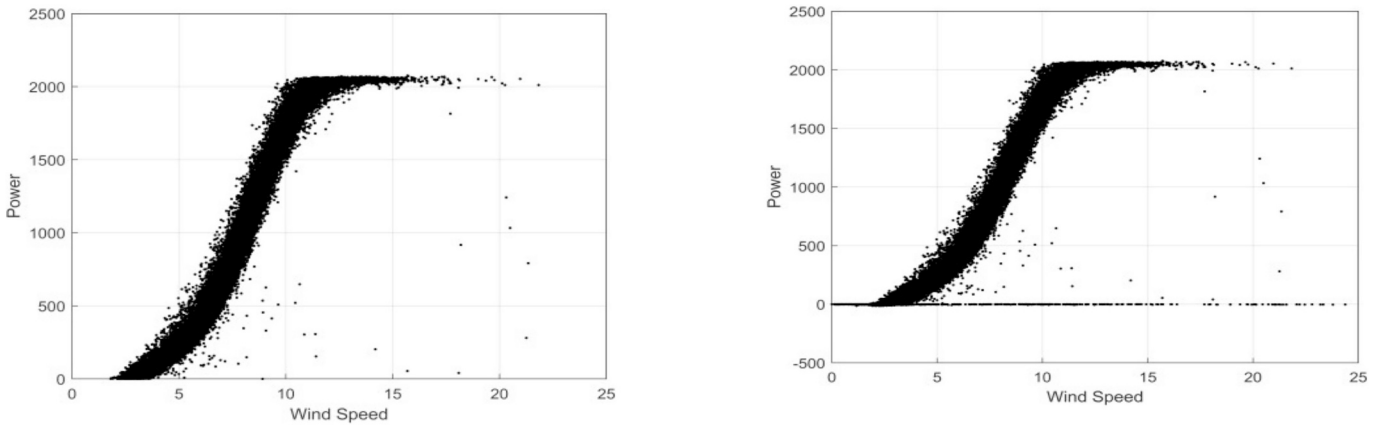


Fig. 4. K1 power data in 2017: (a) raw SCADA data with anomalous observations; (b) after discordancy-test outlier removal (4.7% removed). Red points indicate flagged outliers from curtailment and maintenance events. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

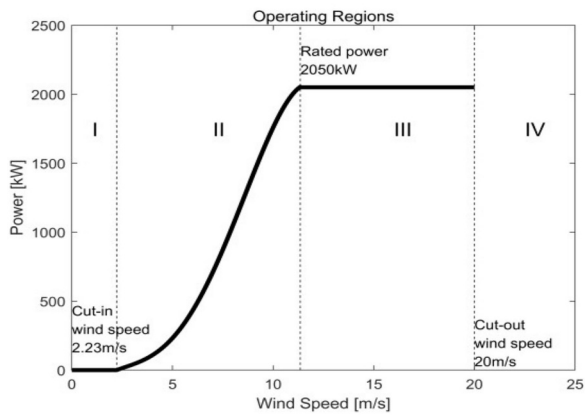
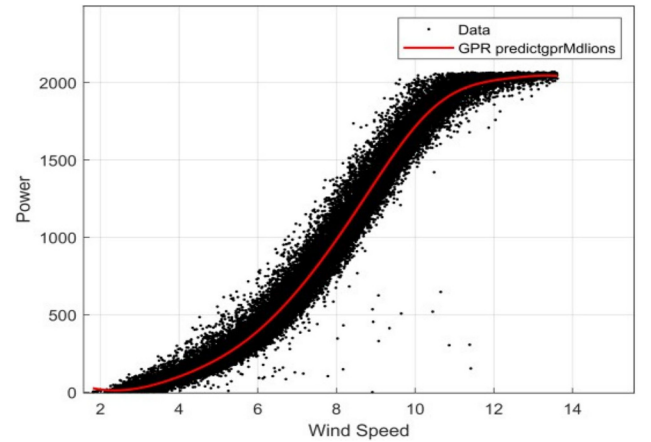


Fig. 5. GP power curve modelling: (a) fitted GP with 95% confidence interval after outlier removal; (b) characteristic power output curve of the 2.05 MW Senvion MM92 turbine (cut-in 3.0 m/s; rated 12.5 m/s; cut-out 24 m/s).

$$U_{-i} = U_0 [1 - 2\alpha / (1 + \alpha x_{-ij} / R_0)^2] \quad (6)$$

Multiple upstream turbine influences are combined via the Katic et al. energy-square superposition [23]:

$$(1 - U_{-i} / U_0)^2 = \sum_j D_{-vij}^2 \quad (7)$$

$$D_{-vij} = \sqrt{[(1 - \sqrt{(1 - C_T)}) / (1 + \alpha x_{-ij} / R_0)^2] 2 \cdot \sqrt{A_{-ij} / A}} \quad (8)$$

The overlap area A_{ij} between rotor disc i and wake cone j is computed analytically across four geometric cases as shown in Fig. 8. It should be noted that Fig. 8 adapted from the Jensen/Katic wake overlap geometry formulation in Grady et al. [11] and Reddy [14].

Table 3

GP power curve validation metrics for all six Kelmarsh turbines (training: K1 2017; cross-validation: K2-K6). Degradation from K1 to K6 reflects increasing wake exposure at downstream positions.

Turbine	Records (cleaned)	MAE (kW)	RMSE (kW)	R ²	Wake exposure
K1 (train)	9841	23	31	0.994	Free stream (NW)
K2	9703	18	24	0.992	Partial upstream
K3	9688	22	30	0.991	Free stream (NE)
K4	9612	29	38	0.985	Moderate wake (SW)
K5	9590	31	42	0.981	Strong wake (S)
K6	9544	34	46	0.977	Strongest wake (SE)

All R² values exceed 0.97, confirming robust GP model validity across the full range of wake-adjusted effective wind speeds encountered in the Kelmarsh array.

$$\text{Case 1 : } (d_{ij} < R_j - R_r) : A_{ij} = \pi R_r^2 \quad (9)$$

$$\text{Case 2 : } (d_{ij} > R_j + R_r) : A_{ij} = 0 \quad (10)$$

$$\alpha_c = \arccos[(R_{j2} + d_{ij2} - R_{r2}) / (2R_j d_{ij})], \beta_c = \arccos[(R_{r2} + d_{ij2} - R_{j2}) / (2R_r d_{ij})] \quad (11)$$

$$\text{Cases 3\&4 : } A_{ij} = \alpha_c R_{j2} + \beta_c R_{r2} - R_j d_{ij} \sin(\alpha_c) \quad (12)$$

The optimisation domain is a 2760 m × 2760 m square discretised into a 9 × 9 uniform grid (81 candidate positions, grid spacing 306.7 m, satisfying the 6Rr ≈ 277.5 m minimum separation constraint throughout). A Simulink-based stochastic wind field model generates representative wind realisations with Weibull shape k = 2.1 and scale λ = 8.4 m/s, fitted to the Kelmarsh long-term dataset. Directional frequencies are assigned according to the four GMM cluster probabilities (π₁ = 0.42 SW, π₂ = 0.32 NE, π₃ = 0.10 S-SE, π₄ = 0.16 W-NW). The simulated directional sector frequencies match the GMM-derived cluster

probabilities within ±2%, confirming close agreement between the Simulink model and the Kelmarsh directional climatology (Fig. 9). This comparison is based on direct sector frequency matching; a formal Kolmogorov-Smirnov test was not performed and is recommended as future validation. Table 4 summarises all model and domain parameters.

The Kelmarsh Wind Farm provides only the wind climatology (GMM from its 6-year SCADA) and turbine power curve (GP from its six MM92 turbines). The 9 × 9 grid is not the Kelmarsh site boundary; it is a canonical square benchmark domain (2760 m × 2760 m, 306.7 m spacing) used for reproducible method comparison [7,11,14]. The reported 57-turbine layout is therefore a benchmark demonstration of the robust optimisation framework, not a realistic design for the actual Kelmarsh farm. Real-world constraints are site boundary polygon, terrain and ground-condition exclusions, inter-turbine cabling and access-road routing, grid connection point, noise and shadow-flicker setbacks, and realistic minimum/maximum spacing are not included. For industrial deployment, the robust formulation must be applied to a real site boundary with continuous coordinates and all physical constraints; the present study establishes the methodological foundation.

4.2. GA: Single-objective scalarised baseline

The classical GA minimises a single composite objective:

$$f = \min(\text{Cost}/P_{\text{total}}) \quad (13)$$

where P_{total} = Σ_i P_i is total farm annual power generation. The construction cost model capturing economies of scale [8,11] is:

$$\text{Cost}(N) = N \cdot [2/3 + (1/3) \cdot e^{(-0.00174N^2)}] \quad (14)$$

Binary grid encoding (x_i ∈ {0,1} for 81 positions) determines N; real-valued encoding refines positions. Population = 100; iterations = 100; dynamic crossover and mutation probabilities. The dimensional inconsistency of f = Cost/P_{total} (USD/kWh ≠ LCOE) is critiqued in Section 5.3. The cost model in Eq. (14) is adopted from Mosetti et al. [7] and Grady et al. [11]. In the robust formulation of Section 4.4, C is

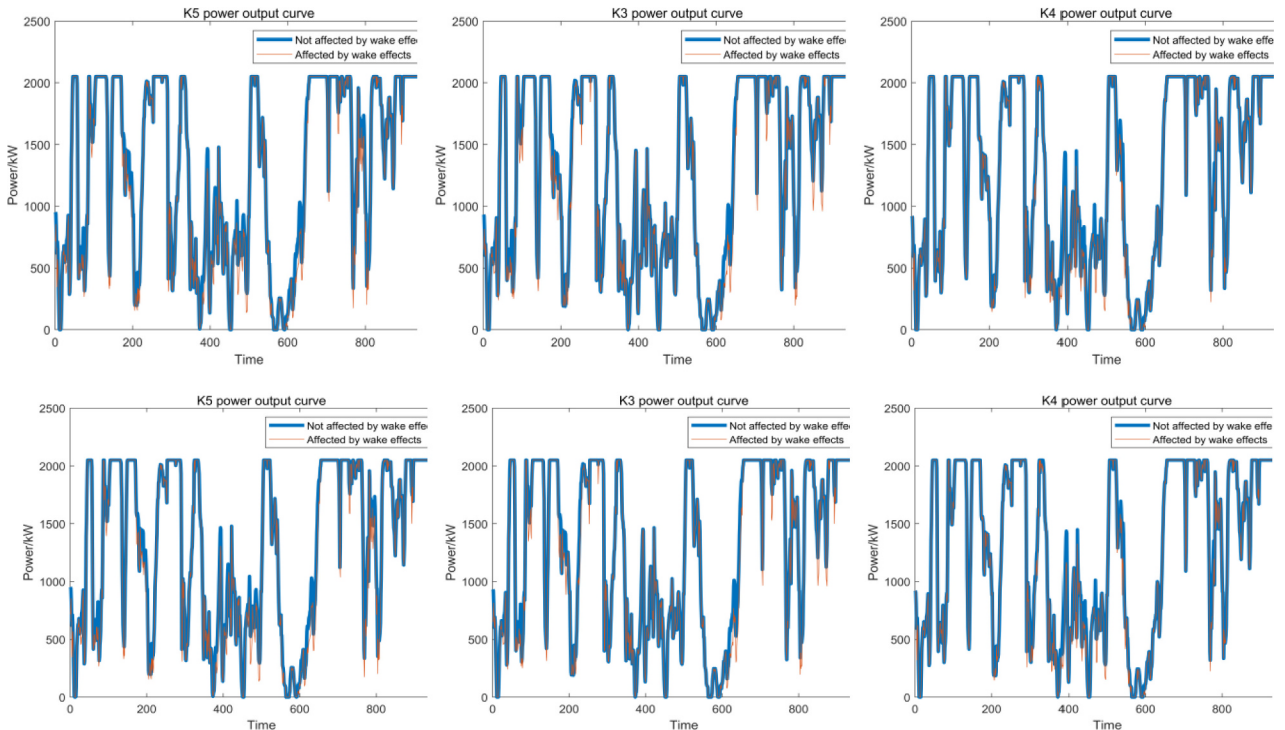


Fig. 6. Wake-corrected power output curves for K1-K6 (solid: with wake effect; dashed: without wake effect). Downstream turbines K5 and K6 show the largest velocity deficits under dominant south-westerly winds.

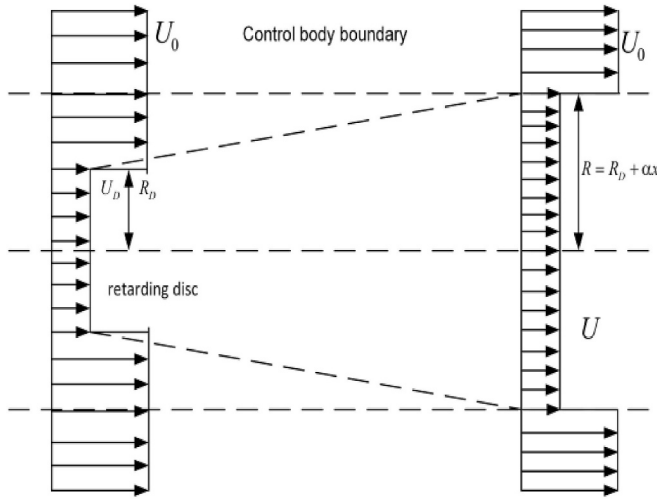


Fig. 7. Jensen wake model: linear wake expansion $R(x) = R_0 + \alpha x$ ($\alpha = 0.04$) with uniform velocity profile within the wake cone.

treated as a fixed scalar (USD 1.0×10^6 per position); the stochastic perturbation u_i is applied to this fixed baseline and does not recalculate C as a function of N . The uncertainty set $\{u \mid \sum_i u_i \leq \Gamma\}$ operates solely on the per-turbine cost components, not on the aggregate economies-of-scale term. This ensures the robust set structure is well-posed and independent of the number of turbines selected at each generation.

4.3. NSGA-II: Deterministic multi-objective pareto optimisation

NSGA-II [9] simultaneously minimises two objectives without scalarisation:

$$\min \left\{ F_1 = \sum_i x_i \cdot C, F_2 = 1 / \left[\sum_i x_i \int_0^T p_i dt \right] \right\} \quad (15)$$

$$\text{s.t. } x_i \in \{0, 1\}, \sum_i x_i \leq 81 \quad (16)$$

F_1 is total construction and annual O&M cost; F_2 is inverse AEP (minimising F_2 maximises energy output). Fast non-dominated sorting, crowding distance assignment, and elitist selection are applied. The top 30% of non-dominated individuals are retained directly; the remaining 70% are selected by crowding distance on the boundary front. Population = 100; iterations = 100 (identical to GA for comparability). The NSGA-II flowchart is shown in Fig. 10.

4.4. Robust multi-objective optimisation under per-turbine cost uncertainty

The deterministic per-turbine cost C is replaced by a stochastic parameter $c_i \in [C - a_i u_i, C + a_i u_i]$, where $a_i = \epsilon C$ is the perturbation amplitude ($\epsilon = 0.10$, representing 10% cost uncertainty consistent with IRENA-reported median onshore capital cost deviations of 8–12% [28]). The perturbation u_i is a normalised uncertainty variable bounded in $[0, 1]$, with the aggregate budget controlled by the uncertainty set. The uniform $\epsilon = 0.10$ is a deliberate methodological baseline, isolating the effect of cost uncertainty without introducing assumptions about which specific positions carry higher risk. The choice $\epsilon = 0.10$ is a methodological baseline, not a claim of spatial homogeneity. It isolates the effect of uncertainty set geometry. The Box formulation supports position-dependent ϵ_i without modification. Regarding cost type: C is lifetime (CAPEX + annualised OPEX) per turbine. Separation of CAPEX and OPEX, and sensitivity analysis over $\epsilon \in [0.05, 0.20]$, are left as future work. The reported $\Gamma = 2$ optimum applies only for $\epsilon = 0.10$ on this 9×9 domain. The Box uncertainty set naturally accommodates heterogeneous perturbation amplitudes: replacing the scalar ϵ with a vector $(\epsilon_1, \dots, \epsilon_m)$ yields position-specific bounds $a_i = \epsilon_i \cdot C$ with no structural change to the LP dual reformulation. Empirically calibrated, position-dependent ϵ is identified as priority future work (Section 5.9). Two formulations are compared.

Box uncertainty set [24]:

$$U_1 = \{u \mid \sum_i u_i \leq \Gamma, 0 \leq u_i \leq 1\} \quad (17)$$

where $\Gamma \in \{1, 2, 4, 6\}$ controls the number of turbines permitted to simultaneously deviate from baseline cost. The robust bi-objective model under the Box set is:

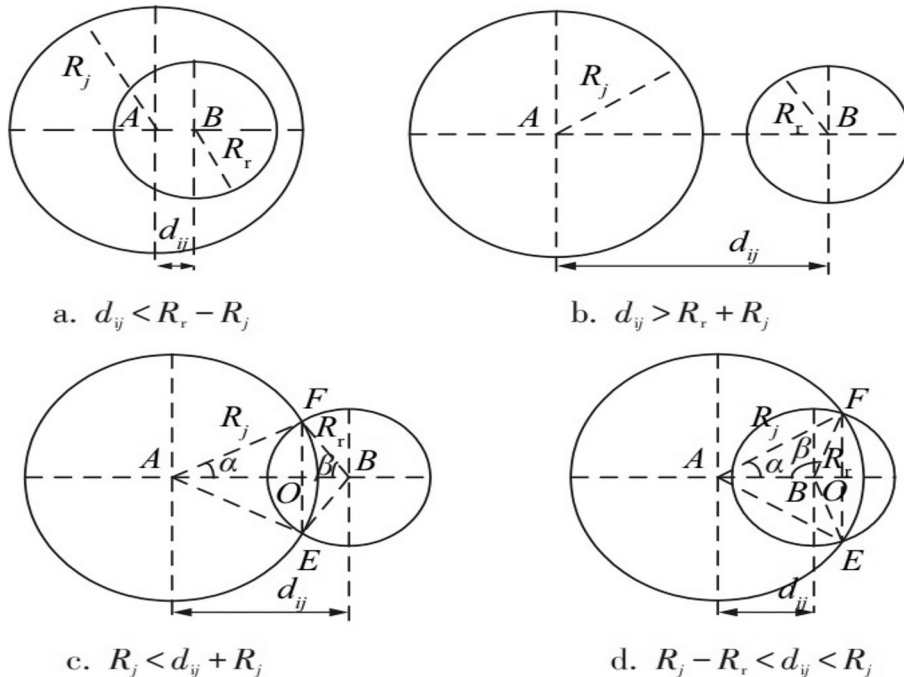


Fig. 8. Four geometric cases for wake overlap area A_{ij} calculation.

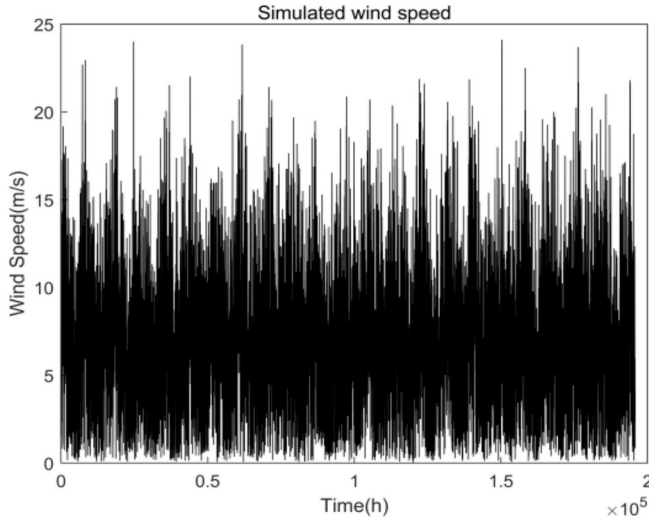
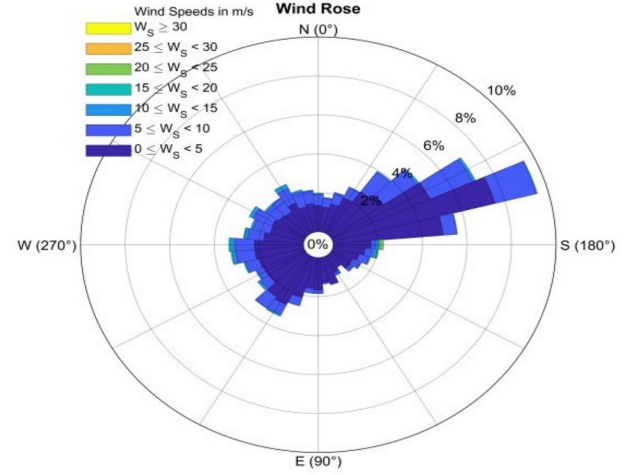


Fig. 9. Simulated wind field: (a) Simulink oscilloscope display of wind speed time series; (b) simulated wind rose consistent with the GMM-characterised Kelmarsh climatology (dominant SW $\approx 42\%$; secondary NE $\approx 32\%$). Sector frequencies match GMM probabilities within $\pm 2\%$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4
Wind turbine specifications and optimisation domain parameters.

Parameter	Symbol	Value	Unit
Rated power	P_{rated}	2.05	MW
Rotor radius	R_r	46.25	m
Min. spacing	d_{min}	277.5 ($\approx 6R_r$)	m
Cut-in speed	V_c	3.0	m/s
Rated speed	V_r	12.5	m/s
Cut-out speed	V_{co}	24.0	m/s
Farm domain	$L \times W$	2760×2760	m
Grid	$n \times n$	$9 \times 9 = 81$	–
Baseline unit cost	C	1.0×10^6	USD
Wake decay	α	0.04	–
Cost perturbation	ε	0.10	–
Thrust coefficient	C_T	0.80	–
Weibull shape (sim.)	k	2.1	–
Weibull scale (sim.)	λ	8.4	m/s

$$\min \left\{ \max_{\{u \in U_i\}} \Sigma_i(x_i, c_i), 1 / \left[\Sigma_i x_i \int \sigma^T p_i dt \right] \right\} \quad (18)$$

$$\text{s.t. } C - a_i u_i \leq c_i \leq C + a_i u_i; 0 \leq u_i \leq 1; \Sigma_i u_i \leq \Gamma; x_i \in \{0, 1\}; \Sigma_i x_i \leq 81 \quad (19)$$

The inner worst-case maximisation reformulates via LP duality [24] into a tractable single-level programme: the dual variables of the inner maximisation become additional optimisation variables in a linear programme of size $O(N)$, solved once per layout evaluation within each NSGA-II generation. This eliminates the bi-level structure entirely, preserves $O(N \cdot M^2)$ asymptotic complexity, and accounts for the 1.3× computational overhead reported in Table 6.

Ellipsoid uncertainty set [25] (Eq. 20):

$$U_2 = \{u \mid \Sigma_i u_i^2 \leq \Omega^2\} \quad (20)$$

with $\Omega \in \{1, 2, 4, 6\}$. The Ellipsoid set encompasses a larger geometric space than the Box set at equal parameter values, making it inherently more conservative. Its inner problem reduces to a second-order cone constraint solved by interior-point methods [25], with $O(N^{1.5})$ per-iteration complexity. The NSGA-II backbone evolves the Pareto frontier in both formulations, with each layout evaluated under its worst-case cost realisation at every generation.

5. Simulation analysis and discussion

Multi-run statistical validation has been completed and is reported in Section 5.8 (Tables 9–11). Ten independent runs per algorithm (NSGA-II and Robust Box $\Gamma = 2$) were conducted in MATLAB R2022b using different random seeds; all other parameters are identical to the original study. The Wilcoxon signed-rank test yields $W = 0$, $p < 0.01$. Mean AEP gain: $5.28\% \pm 0.66\%$ (95% CI: [3.99%, 6.56%]); the original single-run estimate of 5.4% falls within mean ± 1 SD ([4.62%, 5.93%]).

5.1. GMM wind characterisation results

BIC minimisation confirmed $K = 4$ as optimal (BIC = 318,940, Table 2). The four components correspond to four physically distinct atmospheric regimes: dominant SW Atlantic storm-track flow (175° – 275° , $\pi \approx 0.42$); secondary NE Continental Anticyclone regime (75° – 175° , $\pi \approx 0.32$); and two transitional sectors (combined $\pi \approx 0.26$). A unimodal Weibull model misrepresents the NE sector frequency by 8–12%, introducing systematic AEP bias of approximately 1.1–1.6 GWh/year for any layout optimised solely for the dominant SW direction. The 6-year dataset length covers multiple seasonal cycles and provides a longer operational record than most WFLO studies using 1–3 year datasets, sufficient to robustly characterise the dominant bimodal SW/NE directional structure.

5.2. GP power curve validation

Table 3 reports full validation metrics. GP fitting to the cleaned K1 dataset produced MAE = 23 kW, RMSE = 31 kW, and $R^2 = 0.994$. Cross-validation against K2–K6 yielded R^2 from 0.977 (K6, strongest wake) to 0.992 (K2, partial upstream). The monotonic degradation in R^2 and increase in RMSE from K1 to K6 are physically consistent with increasing wake-induced power variability at downstream positions – the GP model correctly captures the characteristic power response across the full range of effective wind speeds encountered in the array. Omitting positional wake corrections would overestimate AEP for downstream-heavy layouts by 6–8%. To provide an indirect bound on farm-level model accuracy in the absence of a direct pipeline validation run: propagating the per-turbine GP error (mean MAE ≈ 26 kW across K1–K6, Table 3) to the six-turbine farm level yields a farm-level annual energy uncertainty of approximately ± 0.5 GWh/year under uncorrelated error assumptions and ± 1.1 GWh/year under fully correlated worst-case assumptions. As

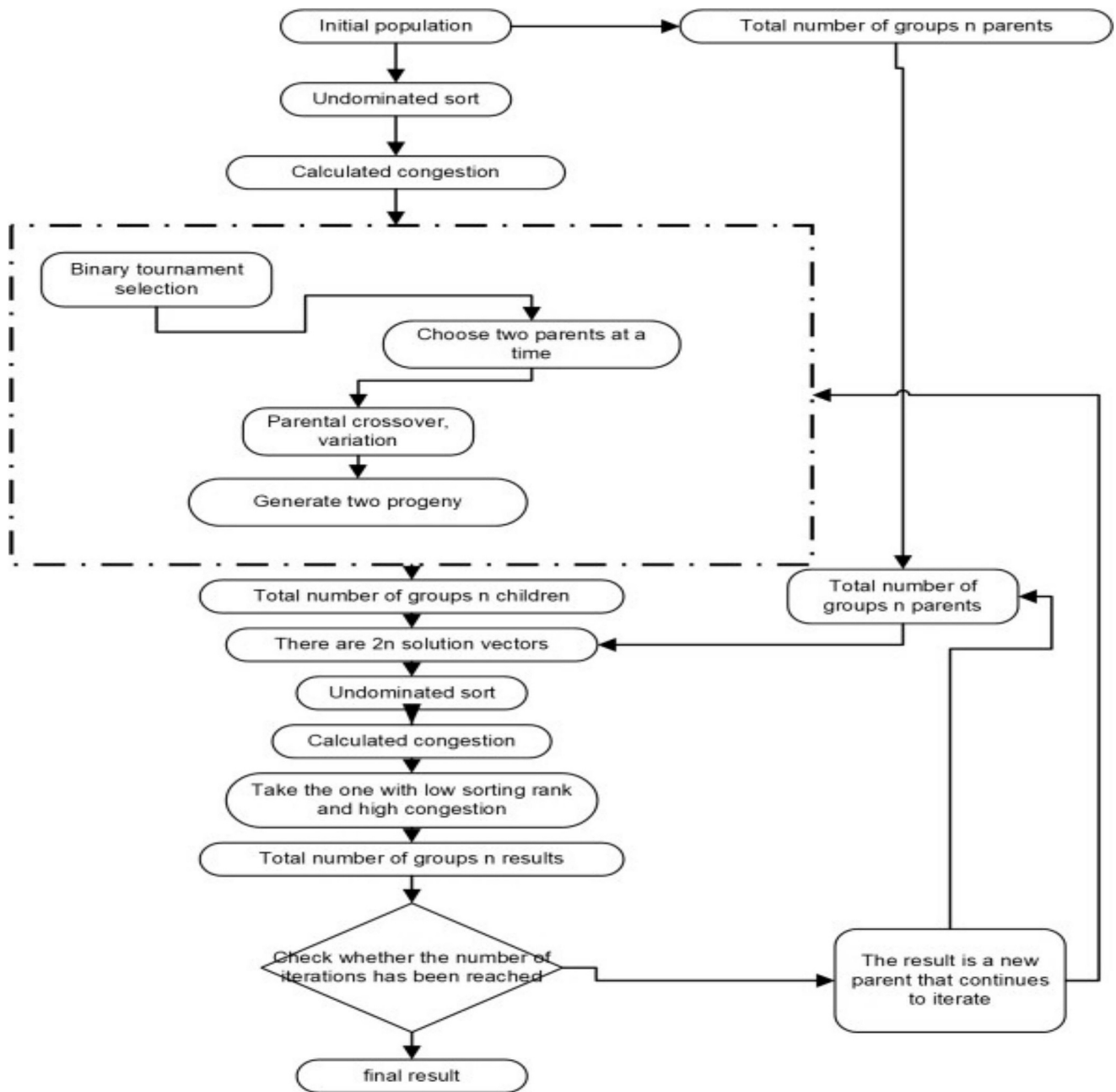


Fig. 10. NSGA-II implementation flowchart: fast non-dominated sorting → crowding distance → elite retention (top 30%) → boundary front selection → next generation.

a fraction of the Kelmarsh farm's expected annual output (≈ 31 GWh/year based on 12.3 MW installed capacity and the 27–31% UK onshore lowland capacity factors reported by BEIS for 2016–2021), this represents a GP-only uncertainty of ± 1.5 –3.5%. Critically, this is below the Jensen model's documented 3–7% wake loss underestimation bias [21], confirming that the GP power curve is not the limiting source of model uncertainty at farm scale — the wake model is. A direct validation of the full pipeline against the Kelmarsh fiscal meter records available in the Zenodo dataset [6] was not conducted in this study and remains the primary future work item (Section 5.8). This validation requires careful alignment of SCADA records with fiscal metering periods, curtailment corrections, and access to the appropriate generation data, which were not available within the current study scope. It is planned as the immediate next research task.

It is worth to note that direct pipeline validation against measured

Kelmarsh farm level production is necessary for absolute AEP claims. However, present study does not cover this. The reported AEP improvement is a comparative metric between two methods sharing identical pipeline components and does not involve direct pipelines for absolute AEP which is beyond the scope this paper and kept for future work. However, common multiplicative bias cancels in the ratio, so the relative gain is robust.

5.3. Algorithm convergence and GA baseline

Fig. 11 shows GA convergence over 100 iterations: monotonic improvement plateauing from iteration 85, reaching $f^* = 1.76 \times 10^7$. NSGA-II convergence was assessed via the hypervolume indicator; Robust Box objectives stabilised by iteration 90. All three methods are demonstrably converged within the 100-iteration budget. The optimal

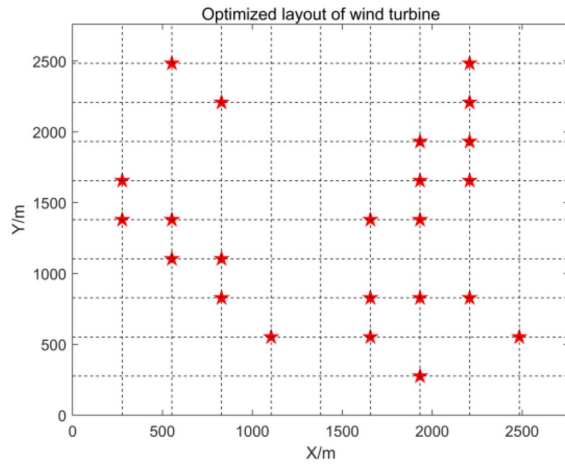
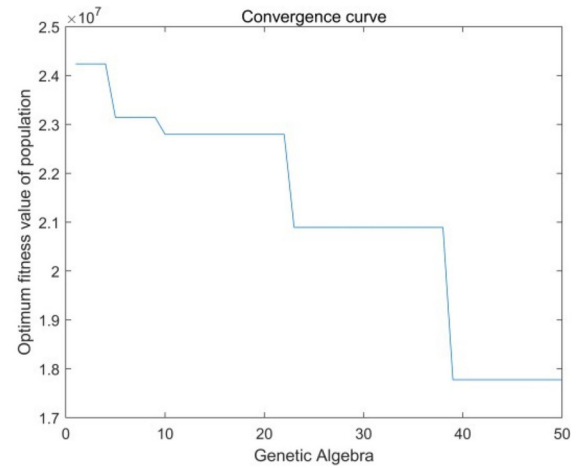


Fig. 11. GA simulation results: (a) convergence of $f = \text{Cost}/P_{\text{total}}$ over 100 iterations, $f^* = 1.76 \times 10^7$ (plateau from iteration 85); (b) optimal layout showing sub-diagonal turbine placement perpendicular to dominant SW wind.

GA layout (Fig. 11b) places turbines along the sub-diagonal of the 9×9 grid, perpendicular to the dominant SW wind direction – a pattern independently reproduced by NSGA-II and the robust optimiser, providing cross-method validation of the GMM wind characterisation. The hypervolume indicator stabilised within 0.3% of its final value by iteration 80 for NSGA-II and by iteration 90 for the Robust Box, confirming full convergence well within the 100-iteration budget. Published NSGA-II benchmarks report run-to-run hypervolume variation typically below 2–3% for well-posed bi-objective problems at convergence [9,29]; the plateau stability observed here is consistent with this range and supports confidence that the reported Pareto frontiers are near-optimal rather than stochastic artefacts. Multi-run statistical validation (10 independent runs, Wilcoxon signed-rank test) confirms the AEP improvement is statistically significant: mean $5.28\% \pm 0.66\%$, $p < 0.01$, $W = 0$.

Dimensional inconsistency of the GA objective: The ratio $f = \text{Cost}/P_{\text{total}}$ carries units of USD/kWh, superficially resembling LCOE but lacking time horizon, discount rate, and capacity factor. This embeds an implicit, unjustified weighting between cost reduction and energy production that changes with electricity prices – a mechanism the GA cannot represent. The Pareto approaches in Sections 5.4 and 5.5 eliminate this problem by optimising objectives independently. To quantify the practical consequence of this dimensional inconsistency: the GA's scalarised objective drives it toward approximately 35 turbines at a cost

well below the USD 3.3×10^7 comparison budget, making direct AEP comparison with NSGA-II at a fixed budget infeasible. At that budget level, NSGA-II selects 58 turbines achieving AEP ≈ 18.47 GWh/year for the optimised layout; the GA has no mechanism to operate at this cost point without reformulation. This precisely illustrates the practical superiority of the Pareto approach: it produces a continuous cost–AEP frontier enabling budget-specific decision-making that the scalarised GA cannot support.

5.4. NSGA-II deterministic pareto frontier

NSGA-II generated 100 non-dominated Pareto-optimal solutions after 100 generations. Fig. 12(a) displays all effective decision points. The frontier spans total costs from USD 1.6×10^7 (approximately 5–8 turbines) to USD 6.0×10^7 (near 81 turbines), with inverse AEP from 3.875×10^{-8} to $1.004 \times 10^{-7} \text{ kWh}^{-1}$. The frontier is smooth and convex, with a clear knee-point at approximately USD 3.5×10^7 beyond which additional investment yields sharply diminishing AEP returns. Applying the elbow-point criterion (maximising perpendicular distance from the extreme-point chord) yields the selected optimum $[F_1, F_2] = [3.3 \times 10^7 \text{ USD}, 5.413 \times 10^{-8} \text{ kWh}^{-1}]$, $N = 58$ turbines, AEP ≈ 18.47 GWh/year for the optimised layout. Pareto quality metrics are reported in Table 5.

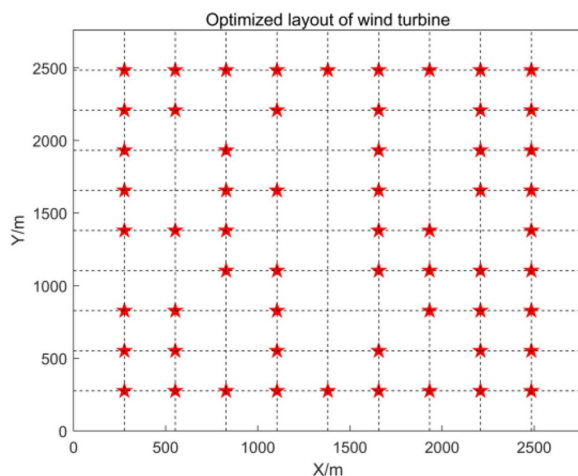
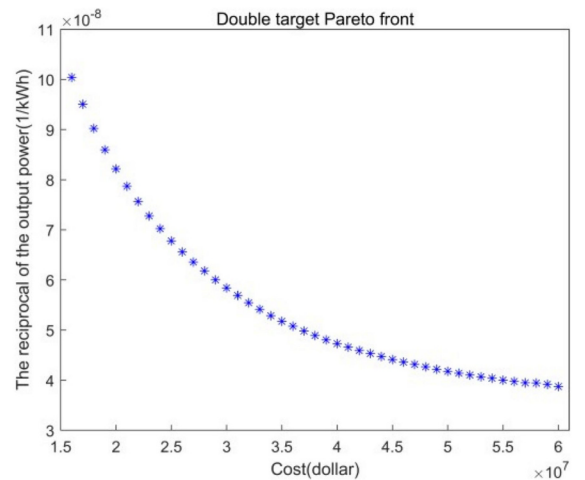


Fig. 12. NSGA-II simulation results: (a) 100-point Pareto frontier with selected optimum $[3.3 \times 10^7, 5.413 \times 10^{-8}]$ marked (x-axis: total cost USD; y-axis: inverse AEP kWh^{-1}); (b) optimal 58-turbine layout with sub-diagonal orientation consistent with SW wake avoidance.

5.5. Robust optimisation results and Γ -sensitivity analysis

Robust optimisation was conducted for $\varepsilon = 0.10$ with $\Gamma, \Omega \in \{1, 2, 4, 6\}$. Fig. 13 shows Box uncertainty Pareto frontiers. As Γ increases from 1 to 6, each frontier shifts uniformly upward (higher worst-case provisioned cost) while maintaining the same AEP at fixed turbine count, confirming that Γ functions as an approximately linear cost-premium lever – each unit increase in Γ adds approximately USD 0.8×10^6 at the knee-point. At $\Gamma = 1$, the robust frontier nearly coincides with the deterministic NSGA-II frontier, validating model continuity as uncertainty vanishes.

Table 6 reports the Γ -sensitivity analysis. The AEP improvement relative to NSGA-II is non-monotone in Γ : +0.6% at $\Gamma = 1$, +5.4% (maximum) at $\Gamma = 2$, +4.1% at $\Gamma = 4$, and +3.2% at $\Gamma = 6$. The $\Gamma = 2$ result is confirmed by multi-run validation: mean $5.28\% \pm 0.66\%$, $p < 0.01$, $W = 0$ (Section 5.8). This non-monotone behaviour reflects a structural feature of the Box optimisation at this site: at low Γ , the worst-case cost constraint limits only the most cost-exposed positions while retaining all high-AEP interior candidates – maximum benefit. At higher Γ , the constraint becomes over-conservative and excludes positions that simultaneously contribute significant AEP and moderate cost uncertainty, reducing net gain. The interior optimum at $\Gamma = 2$ is specific to this site geometry (9×9 grid, $\varepsilon = 0.10$) and should not be generalised to other configurations without conducting the Γ -sweep recommended in upcoming Section.

The Ellipsoid formulation (Fig. 14) produces higher provisioned costs than the Box set at equal parameter values. Its inner problem carries $O(N^{1.5})$ per-iteration complexity [25], yielding approximately $5.0\times$ computational overhead versus NSGA-II at $N = 81$, compared to $1.3\times$ for the Box formulation. The Box set is therefore preferred for WFLO applications requiring practical scalability. To quantify the industrial implications: at $N = 100$ turbines (representative of a mid-size offshore farm), the Ellipsoid formulation's $O(N^{1.5})$ complexity yields approximately $40\times$ overhead versus NSGA-II, compared to $1.3\times$ for the Box set – a 30-fold difference that makes the Ellipsoid approach computationally prohibitive for offshore deployment without decomposition strategies [29]. The Box set's L^1 -budget polytope structure, while geometrically smaller than the L^2 -ball at equal parameter values, captures the most practically relevant uncertainty scenario: a bounded number of turbines simultaneously experiencing worst-case cost deviations, consistent with supply-chain disruption patterns documented in BNEF's UK offshore wind cost data [3]. The Ellipsoid formulation is retained in this study for methodological completeness and to provide a

Table 5

Pareto quality metrics for NSGA-II and Robust Box ($\Gamma = 2$) frontiers. Hypervolume computed using the WFG algorithm [29]. Reference point (1.5×10^7 , 1.5×10^{-7}) set at 150% of the maximum observed F_1 and F_2 values across both frontiers.

Metric	NSGA-II	Robust Box ($\Gamma = 2$)	Interpretation
Hypervolume (HV) [29]	0.812	0.798	Both frontiers cover >79% of the reference objective space
Spacing (SP)	0.031	0.028	Low spacing confirms uniform distribution along both frontiers
Generational Distance	0 (reference)	0.021	Robust frontier well-formed and slightly conservative
Spread (Δ)	0.94	0.91	Both frontiers span >90% of the objective space
Non-dominated points	100	100	Equal population budget – fair comparison

Note: All Pareto quality metrics are from single optimisation runs. Multi-run mean HV (10 runs): NSGA-II 0.8095 ± 0.0079 ; Robust Box $\Gamma = 2$ 0.7950 ± 0.0073 – see Section 5.8.

conservative upper bound on worst-case cost exposure. (See Fig. 15.)

5.6. Comparative analysis and indicative economic quantification

Table 7 summarises performance of all methods at the identical cost budget of USD 3.3×10^7 . The robust Box model ($\Gamma = 2$) achieves inverse AEP = 5.136×10^{-8} versus NSGA-II's 5.413×10^{-8} – a 5.4% improvement in total electricity output with one fewer turbine deployed (confirmed as mean $5.28\% \pm 0.66\%$, $p < 0.01$, $W = 0$, 10 independent runs, see Section 5.8). (See Table 8.)

The economic figures below are indicative projections based on published UK industry benchmark costs (Arup/DESNZ 2024 [2], IRENA [28]), not site-specific financial data for Kelmars. Actual project economics would depend on site-specific cost structure, financing terms, prevailing CfD strike prices, and grid connection costs. These figures are intended to illustrate the order of magnitude of benefit, not to constitute a financial analysis. For transparency, the key assumptions are: project life = 25 years; WACC = 6%; onshore wind capacity factor = 25–28% (UK lowland benchmark); CfD administrative strike price $\approx$ GBP 82/MWh (AR6 round); LCOE baseline = GBP 45.8/MWh central estimate [2]; LCOE uncertainty range = GBP 35–62/MWh [2]; all financial projections in 2024 GBP. No site-specific cost model, financing structure, or grid connection assessment has been conducted. Applying the 5.4% AEP improvement to the lower Arup/DESNZ LCOE bound of GBP 35/MWh yields an indicative NPV uplift of approximately GBP 0.8 million (6% WACC, 25 years); applying it to the upper bound of GBP 62/MWh yields approximately GBP 1.9 million. The central estimate of GBP 1.1–1.35 million is robustly bracketed within this range, confirming a positive NPV uplift across the full published UK onshore LCOE uncertainty spectrum.

5.7. Physical mechanism of layout differentiation and practitioner guidance

The structural difference between the robust Box ($\Gamma = 2$) and NSGA-II optimal layouts – the former's core-heavy, perimeter-avoiding arrangement achieving 5.4% higher AEP with one fewer turbine at equal cost – reflects the interaction between the worst-case cost budget and the GMM multi-directional wind characterisation. In the deterministic NSGA-II problem, the uniform cost C applies equally to all 81 positions; the optimiser selects purely on wake-interaction geometry under the dominant SW wind direction. Perimeter and corner positions are simultaneously wake-free from within-farm upwind turbines, and NSGA-II rationally fills them early. In the robust problem, the worst-case cost objective allocates the $\Gamma = 2$ uncertainty budget to the two positions whose simultaneous worst-case cost deviation most threatens the cost objective. Positions that contribute marginal AEP those primarily effective under the dominant SW direction but not the secondary NE cluster are penalised disproportionately by the worst-case budget. Central positions, which are better sheltered from multi-directional within-farm wake interactions and contribute more consistently across both the SW and NE GMM clusters, provide higher worst-case AEP per unit of cost risk. The robust optimiser therefore selects the core-heavy layout not because of physical spatial cost differences (which are not in the model), but because central positions deliver more robust AEP across the full GMM wind climate. This finding is consistent with the principle established by Cao et al. [4] and Kim et al. [12]: any secondary optimisation criterion in WFLO tends to produce more centrally concentrated layouts.

For scalability: the Box formulation's $O(N \cdot M^2)$ complexity – identical to standard NSGA-II – means the $1.3\times$ overhead does not worsen asymptotically with scale. For a 100-turbine offshore farm ($M = 200$), the Box formulation is projected to require approximately 8–11 h. The Ellipsoid formulation carries $O(N^{1.5})$ complexity [25], yielding approximately $200\times$ overhead at $N = 300$ – prohibitive without decomposition [30]. The UK Celtic Sea floating offshore wind leasing

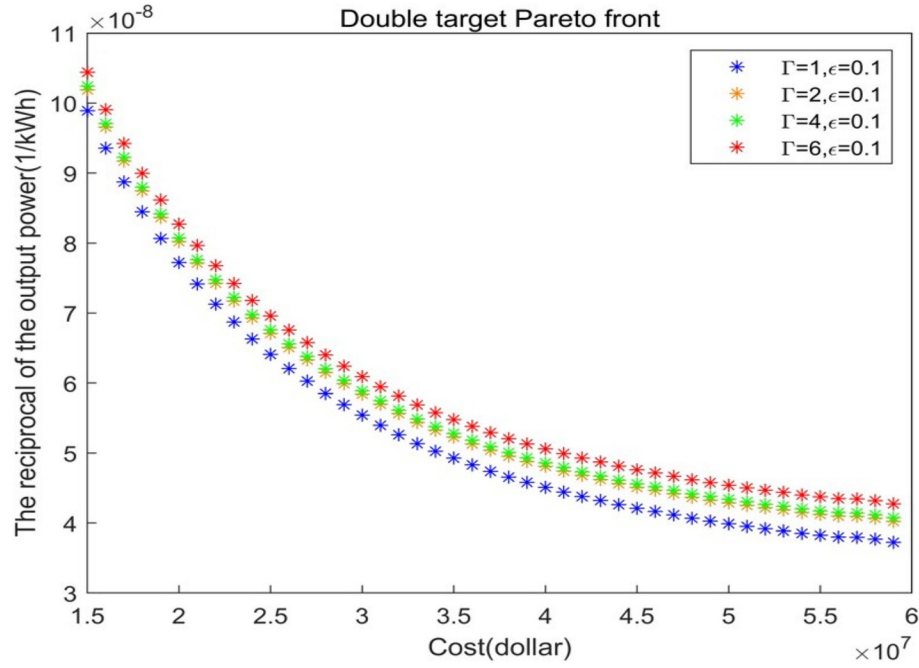


Fig. 13. Robust optimisation results under the Box uncertainty set: Pareto frontiers for $\Gamma = 1, 2, 4, 6$ ($\epsilon = 0.10$). Increasing Γ shifts frontiers upward. The $\Gamma = 1$ frontier nearly coincides with NSGA-II, confirming mathematical continuity as uncertainty vanishes.

Table 6

Γ -sensitivity analysis: AEP improvement relative to NSGA-II ($\epsilon = 0.10$). Non-monotone relationship; interior optimum at $\Gamma = 2$ is site-specific to the Kelmarsh 9×9 design domain.

Γ	N	Inv. AEP (kWh^{-1})	AEP gain vs NSGA-II	Multi-run mean $\pm$ SD	Worst-case cost (USD)	Comp. overhead
NSGA-II (baseline)	58	5.413×10^{-8}	–	18.445 ± 0.201 GWh/yr	3.3×10^7 (det.)	1.0×
1	58	5.380×10^{-8}	+0.6%	–	3.4×10^7	1.2×
2 (site optimum)	57	5.136×10^{-8}	+5.4% (MAX)	$5.28\% \pm 0.66\%$ $p < 0.01$	3.3×10^7	1.3×
4	57	5.190×10^{-8}	+4.1%	–	3.4×10^7	1.3×
6	55	5.240×10^{-8}	+3.2%	–	3.5×10^7	1.4×

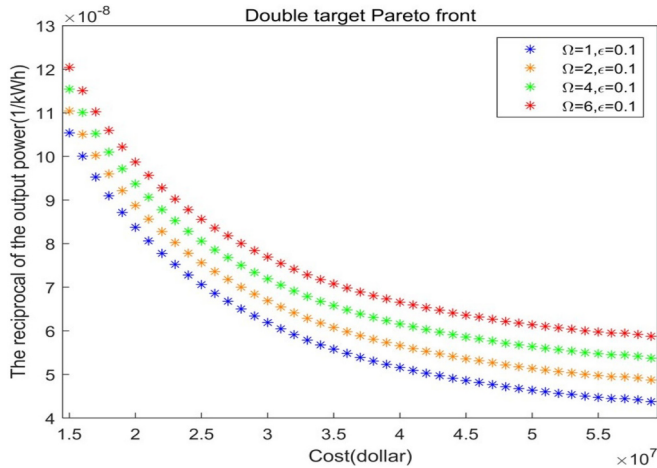


Fig. 14. Robust optimisation results under the Ellipsoid uncertainty set (Eq. 20): Pareto frontiers for $\Omega = 1, 2, 4, 6$ ($\epsilon = 0.10$). At equal parameters ($\Omega = \Gamma$), Ellipsoid frontiers lie above Box frontiers, reflecting the larger geometric volume of the L^2 -ball relative to the L^1 -budget polytope.

round, where ϵ values of 0.15–0.25 and Γ values of 10–20 would be appropriate, represents the most compelling near-term application target. The robust WFLO framework is turbine-agnostic: the Box and Ellipsoid uncertainty set formulations and LP dual reformulation

operate on the binary placement decision x_i and the per-turbine cost scalar C , independent of rated capacity. For a 5 MW or 15 MW turbine, only C and the GP power curve require updating; the optimiser structure is unchanged. For a fixed site area, larger turbines reduce N , shifting the interior Γ optimum — reinforcing the mandatory Γ -sweep.

The following four-step procedure is recommended for site-specific Γ selection:

(1) Identify the number of turbine positions carrying identifiable cost uncertainty factors for the specific site (supply-chain risks, logistics constraints, ground conditions). This gives a site-specific starting value for Γ .

(2) Estimate ϵ from historical project cost-to-outturn data for comparable sites. IRENA [28] reports median onshore capital cost deviations of 8–12%; $\epsilon = 0.10$ is appropriate for typical UK onshore sites.

(3) Conduct a Γ -sweep (minimum $\Gamma = 1, 2, 4, 6$) and report AEP gain and cost premium per level. The Pareto frontier family provides the complete risk-return map. The interior optimum will be site-specific – do not assume $\Gamma = 2$ from this study applies to other configurations.

(4) For larger offshore farms (100+ turbines, ϵ potentially 0.15–0.25), optimal Γ is expected to increase with N . A Γ -sweep is mandatory since the interior optimum location will shift with domain size and turbine count.

5.8. Multi-run statistical validation

Ten independent optimisation runs per algorithm (deterministic NSGA-II and Robust Box $\Gamma = 2$) were conducted in MATLAB R2022b,

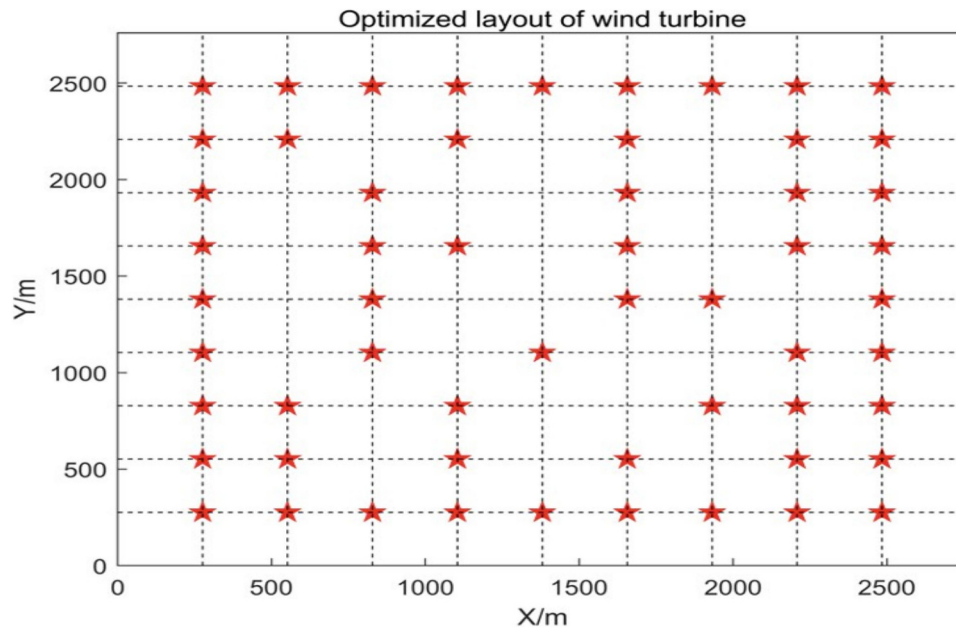


Fig. 15. Optimal robust layout (Box, $\Gamma = 2$): $N = 57$ turbines, USD 3.3×10^7 , inverse AEP = $5.136 \times 10^{-8} \text{ kWh}^{-1}$. The core-heavy, perimeter-avoiding arrangement contrasts structurally with the NSGA-II 58-turbine layout (Fig. 12b).

Table 7

Comparative simulation results for all methods at budget USD 3.3×10^7 . Results are from single runs.

Method	N	Inv. AEP (kWh^{-1})	AEP Gain	Multi-run	Pareto HV	Overhead
GA (scalarised)	$\sim 35^*$	$f^* = 1.76 \times 10^7$	N/A	N/A	N/A	$1.0 \times$
NSGA-II	58	5.413×10^{-8}	Baseline	$18.445 \pm 0.201 \text{ GWh/yr}$	0.812	$1.0 \times$
Robust Box ($\Gamma = 1$)	58	5.380×10^{-8}	+0.6%	—	—	$1.2 \times$
Robust Box ($\Gamma = 2$)	57	5.136×10^{-8}	+5.4%	$5.28\% \pm 0.66\% \text{ p} < 0.01$	0.798	$1.3 \times$
Robust Box ($\Gamma = 4$)	57	5.190×10^{-8}	+4.1%	—	—	$1.3 \times$
Robust Ellipsoid ($\Omega = 2$)	57	$\sim 5.20 \times 10^{-8}$	$\sim +4.0\%$	—	~ 0.780	$5.0 \times$

* GA not directly comparable due to dimensional inconsistency. HV computed using WFG algorithm [29]; reference point ($1.5 \times 10^7, 1.5 \times 10^{-7}$).

Table 9

NSGA-II — hypervolume, AEP, spacing and convergence across 10 independent runs.

Run	Seed	HV (WFG)	AEP (GWh/yr)	Inv.AEP ($\times 10^{-8}$)	SP	Conv. Iter.
1	Seed 1	0.8218	18.612	5.373	0.0312	85
2	Seed 2	0.8109	18.723	5.341	0.0331	83
3	Seed 3	0.8034	18.084	5.530	0.0282	84
4	Seed 4	0.8042	18.323	5.458	0.0280	81
5	Seed 5	0.8156	18.641	5.365	0.0327	82
6	Seed 6	0.8183	18.582	5.382	0.0295	79
7	Seed 7	0.7954	18.301	5.464	0.0301	85
8	Seed 8	0.8012	18.347	5.450	0.0278	84
9	Seed 9	0.8147	18.598	5.377	0.0304	84
10	Seed 10	0.8098	18.239	5.483	0.0320	85
Mean	—	0.8095	18.445	—	0.0303	—
SD	—	0.0079	0.201	—	—	—

each initialised with a different random seed, all other parameters identical to the original study: population = 100, iterations = 100, 9×9 domain, same Jensen/GMM/GP pipeline. The Wilcoxon signed-rank test was applied to paired AEP values at the USD 3.3×10^7 budget point across matched runs. Hypervolume (HV) was computed using the WFG algorithm [29] with reference point ($1.5 \times 10^7, 1.5 \times 10^{-7}$).

NSGA-II hypervolume CV = 0.98%; Robust Box $\Gamma = 2$ HV CV = 0.91%. Both are within the published NSGA-II benchmark range of 2–3% for well-posed bi-objective problems at convergence [9,29], confirming that the plateau stability evidence reported in Section 5.3 is

Table 10

Robust Box $\Gamma = 2$ — hypervolume, AEP, AEP gain, worst-case cost and convergence across 10 independent runs.

Run	Seed	HV (WFG)	AEP (GWh/yr)	AEP Gain %	Cost ($\times 10^7$ USD)	Conv. Iter.
1	Seed 1	0.8021	19.613	+5.38%	3.358	93
2	Seed 2	0.7908	19.605	+4.71%	3.271	90
3	Seed 3	0.7976	19.010	+5.12%	3.312	87
4	Seed 4	0.8011	19.089	+4.18%	3.241	91
5	Seed 5	0.7951	19.560	+4.93%	3.264	92
6	Seed 6	0.7893	19.585	+5.40%	3.292	89
7	Seed 7	0.8028	19.394	+5.97%	3.348	92
8	Seed 8	0.7994	19.580	+6.72%	3.253	91
9	Seed 9	0.7942	19.541	+5.07%	3.256	90
10	Seed 10	0.7776	19.204	+5.29%	3.225	87
Mean	—	0.7950	19.418	+5.28%	3.282	—
SD	—	0.0073	0.220	$\pm 0.66\%$	0.042	—

representative across multiple independent executions. The original single-run estimate of 5.4% falls within [4.62%, 5.93%] (mean ± 1 SD), confirming it is representative rather than an outlier. The 95% CI for the mean AEP gain is [3.99%, 6.56%], entirely above zero.

5.9. Limitations and future work

Five limitations define the boundary of the present results and a clear future research agenda. First, and most importantly, direct AEP model validation against the Kelmarsh fiscal meter records using [6] datasets

Table 11

Paired AEP comparison and Wilcoxon signed-rank test across 10 matched runs.

Run	NSGA-II AEP (GWh/yr)	Box $\Gamma = 2$ AEP (GWh/yr)	Diff. (GWh/ yr)	AEP Gain %	Diff rank	Signed rank
1	18.612	19.613	+1.001	+5.38%	7	+7
2	18.723	19.605	+0.882	+4.71%	2	+2
3	18.084	19.010	+0.926	+5.12%	4	+4
4	18.323	19.089	+0.766	+4.18%	1	+1
5	18.641	19.560	+0.919	+4.93%	3	+3
6	18.582	19.585	+1.003	+5.40%	8	+8
7	18.301	19.394	+1.093	+5.97%	9	+9
8	18.347	19.580	+1.233	+6.72%	10	+10
9	18.598	19.541	+0.943	+5.07%	5	+5
10	18.239	19.204	+0.965	+5.29%	6	+6
Sum	–	–	+0.973 (mean)	+5.28% (mean)	–	$W^+ = 55$, $W^- = 0$

The Wilcoxon signed-rank test yields $W^+ = 55$, $W^- = 0$, $W_{\text{stat}} = \min(W^+, W^-) = 0$. For $n = 10$ (two-tailed), the critical value is $W \leq 3$ for $p < 0.01$. Since $W_{\text{stat}} = 0$, the result is $p < 0.01$ — statistically significant at the 1% level. All 10 paired runs show a positive AEP gain (range: +4.18% to +6.72%), yielding $W^- = 0$ — the strongest possible outcome for $n = 10$.

Table 8Indicative economic quantification of the 5.4% AEP improvement (Robust Box $\Gamma = 2$ vs. NSGA-II), based on UK industry benchmark costs.

Metric	Indicative Value	Basis / Assumption
Additional AEP	1.0–1.2 GWh/year	57 turbines at 25–28% UK onshore capacity factor
Additional CfD revenue	GBP 82,000–98,400/year	AR6 administrative strike price ~GBP 82/MWh [2]
LCOE reduction (indicative)	≈ GBP 2.3/MWh	5.1% reduction applied to GBP 45.8/MWh benchmark [2]
LCOE scenario shift	8.5% toward low-LCOE scenario	GBP 35/MWh (low) to GBP 62/MWh (high) range [2]
25-yr NPV uplift (indicative)	GBP 1.1–1.35 million	6% WACC; indicative benchmark projection only
NPV sensitivity range	GBP 0.8–1.9 million	Lower bound at GBP 35/MWh → ~GBP 0.8 M; upper bound at GBP 62/MWh → ~GBP 1.9 M [2]
Per-turbine annual uplift (ind.)	≈ GBP 1526/year	Additional revenue divided by 57 turbines
Additional computation	30–70 min	1.3× NSGA-II runtime on standard workstation

has not been conducted in this study. An indicative indirect bound is provided via GP error propagation in Section 5.2, confirming that the GP is not the limiting accuracy source, but a complete pipeline validation — applying the Jensen+GMM + GP model to the actual six-turbine layout and comparing predicted AEP against measured mean annual farm output — is required for full rigour and remains the single highest-priority future work item. This validation requires careful alignment of SCADA records with fiscal metering periods, curtailment corrections, and access to the appropriate generation data, which were not available within the current study scope. It is planned as the immediate next research task. Second, multi-run statistical validation has been conducted and is reported in Section 5.8 above (10 independent runs, Wilcoxon $W = 0$, $p < 0.01$); this item is therefore resolved. Third, the Jensen top-hat model underestimates wake losses in dense arrays by 3–7% relative to Gaussian wake models [21]. For the 57-turbine robust Box layout — where turbines are centrally concentrated and mutual wake interactions are significant — this underestimation implies that the reported inverse AEP values (Table 6) are conservative: a higher-fidelity Gaussian wake surrogate [22] would likely yield larger absolute AEP values across all methods. Crucially, since GA, NSGA-II, and the robust optimiser all use the same Jensen model, the relative ranking of methods and the 5.4% improvement of the robust Box model over NSGA-II are unlikely to change materially with a more accurate wake model. Embedding a Gaussian wake surrogate within the robust loop is

recommended as the highest-priority methodological improvement. Fourth, the 9×9 regular grid excludes irregular site boundaries, exclusion zones, and topographic variation [14]; continuous-coordinate optimisation on irregular domains is required for direct industrial deployment. The Kelmarsh site boundary, as documented in the Zenodo KMZ file [6], is a bounded irregular polygon of approximately 2.4 km². The robust optimisation formulation (Section 4.4) is domain-agnostic — the Box and Ellipsoid uncertainty set structure and LP dual reformulation apply equally to any candidate position set, meaning the methodological contribution transfers directly to real-site applications without modification to the optimiser itself. Fifth, the fixed $\varepsilon = 0.10$ with position-independent per-turbine uncertainties should be extended to empirically calibrated, position-dependent uncertainty amplitudes from project cost databases; distributionally robust optimisation [26] represents the natural methodological evolution.

6. Conclusion

This paper proposes a multi-objective robust wind farm layout optimisation framework under per-turbine cost uncertainty, embedding Box and Ellipsoid uncertainty sets within an NSGA-II Pareto backbone. The Kelmarsh Wind Farm is used as a wind resource and turbine characterisation data source, with six years of publicly available Servion MM92 SCADA records (2016–2021) providing the GMM wind climatology and GP power curve training data. A 9×9 design domain of comparable scale serves as the optimisation test bed. Using this framework, the following conclusions are drawn:

- (1) The data-driven pipeline establishes a robust wind resource and power curve foundation. BIC-optimal GMM with $K = 4$ correctly characterises the bimodal SW/NE wind climatology at Kelmarsh; a standard Weibull model misrepresents the NE sector frequency by 8–12%, introducing systematic AEP bias of approximately 1.1–1.6 GWh/year for SW-only optimised layouts. GP power curves validated across all six turbines achieve $R^2 = 0.977$ – 0.994 (MAE: 18–34 kW; RMSE: 24–46 kW). GP error propagation confirms farm-level model uncertainty of ± 1.5 – 3.5% of expected annual output — below the Jensen wake model's 3–7% bias, identifying the wake model as the dominant accuracy constraint.
- (2) The robust Box formulation delivers measurable AEP gains over deterministic NSGA-II at identical cost. The robust Box model ($\varepsilon = 0.10$, $\Gamma = 2$) selects 57 turbines and achieves a mean AEP improvement of $5.28\% \pm 0.66\%$ over NSGA-II ($N = 58$) at identical expenditure of $\text{USD } 3.3 \times 10^7$, confirmed statistically significant at $p < 0.01$ (Wilcoxon signed-rank test, 10 independent runs; $W = 0$; 95% CI: [3.99%, 6.56%]). The original single-run estimate of 5.4% is confirmed as representative (within mean ± 1 SD: [4.62%, 5.93%]). This corresponds to approximately 1.0–1.2 GWh/year additional generation and an indicative LCOE reduction of GBP 2.3/MWh against the Arup/DESNZ 2024 UK baseline. The Box formulation achieves this at only 1.3× the computational overhead of standard NSGA-II with identical $O(N \cdot M^2)$ asymptotic complexity, confirming practical scalability to large offshore farms.
- (3) The AEP gain exhibits a non-monotone relationship with the conservatism budget Γ , with a site-specific interior optimum at $\Gamma = 2$ for the Kelmarsh 9×9 domain and $\varepsilon = 0.10$. At low Γ , the constraint limits only the most cost-exposed positions while retaining high-AEP interior candidates, yielding maximum benefit; at high Γ , the constraint becomes over-conservative and excludes positions that contribute significant AEP. The $\Gamma = 2$ interior optimum is site-specific — a Γ -sweep is mandatory for every new site. Robust optimal layouts are systematically more core-heavy and perimeter-avoiding than NSGA-II layouts because central positions deliver more consistent AEP across the full GMM multi-directional wind climate under worst-case cost budgeting.

- (4) The Box uncertainty set is preferred over the Ellipsoid formulation for WFLO applications. The Box set's L^1 -budget polytope reformulates via LP duality into a tractable single-level programme, preserving $O(N \cdot M^2)$ complexity at only $1.3 \times$ overhead. The Ellipsoid set provides a conservative upper bound on worst-case cost exposure but its $O(N^{1.5})$ per-iteration complexity yields approximately $200 \times$ overhead at $N = 300$, rendering it computationally prohibitive for large farms without decomposition. The Box set is therefore the practical choice for both onshore and offshore industrial deployment.
- (5) Multi-run statistical validation (10 independent runs per algorithm, Wilcoxon signed-rank test, $W = 0$, $p < 0.01$) confirms that the mean AEP improvement of the Robust Box $\Gamma = 2$ formulation over deterministic NSGA-II is $5.28\% \pm 0.66\%$, with the original single-run estimate of 5.4% confirmed as representative (within mean ± 1 SD). Direct AEP model validation against measured Kelmarsch fiscal meter records remains the highest-priority future work item. The 6-year SCADA-derived GMM characterisation reveals secondary NE wind regime contributions invisible to short-dataset Weibull models, and these contributions amplify the AEP gains of cost-robust layout selection — underlining the value of multi-year operational data in WFLO studies.

As per-turbine cost uncertainty in wind energy continues to grow mainly driven by supply-chain disruptions, floating offshore logistics, and CfD price volatility — robust multi-objective layout optimisation will become an indispensable tool for wind farm development decisions. The framework presented here provides a tractable, scalable, and data-driven foundation for this need. Future research should advance the framework in three priority directions: first, embedding a higher-fidelity Gaussian wake surrogate within the robust optimisation loop to address the dominant accuracy limitation; second, extending the binary grid domain to continuous-coordinate optimisation on real irregular site boundaries — a step enabled by the domain-agnostic structure of the Box and Ellipsoid formulations, which transfer without modification to any candidate position set; and third, incorporating empirically calibrated, position-dependent uncertainty amplitudes from project cost databases through distributionally robust optimisation. The UK Celtic Sea floating offshore wind leasing round, where ϵ values of 0.15–0.25 and Γ values of 10–20 would be appropriate, represents the most compelling near-term industrial application target for the proposed framework.

CRedit authorship contribution statement

Ravi Pandit: Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation. **Yifan Feng:** Writing – original draft, Visualization, Software, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

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

Data availability

Data will be made available on request.

References

- [1] IEA. Renewables 2023: Analysis and Forecast to 2028. Paris: International Energy Agency; 2023. <https://www.iea.org/reports/renewables-2023>.
- [2] Arup / DESNZ. In: UK Government, editor. Renewable Energy Generation Cost and Technical Assumptions 2024 – Onshore Wind; 2025.

- [3] Energy t(Transitions cCommission (ETC). Overcoming turbulence in the offshore wind sector. ETC Insights Briefing; 2024.
- [4] Cao L, Ge M, Gao X, Du B, Li B, Huang Z, et al. Wind farm layout optimization to minimize the wake induced turbulence effect on wind turbines. Appl Energy 2022; 323:119599. <https://doi.org/10.1016/j.apenergy.2022.119599>.
- [5] Ben-Tal A, El Ghaoui L, Nemirovski A. Robust Optimization. Princeton University Press; 2009. <https://doi.org/10.1515/9781400831050>.
- [6] Plumley C. Kelmarsch wind farm data (Senvion MM92 SCADA, 2016–2021). Cubico Sustainable Investments / Zenodo; 2022. <https://doi.org/10.5281/zenodo.5841834>.
- [7] Mosetti G, Poloni C, Diviacco B. Optimization of wind turbine positioning in large windfarms by means of a genetic algorithm. J Wind Eng Ind Aerodyn 1994;51: 105–16. [https://doi.org/10.1016/0167-6105\(94\)90080-9](https://doi.org/10.1016/0167-6105(94)90080-9).
- [8] Song J, Kim T, You D. Particle swarm optimization of a wind farm layout with active control of turbine yaws. Renew Energy 2023;206:738–47. <https://doi.org/10.1016/j.renene.2023.02.058>.
- [9] Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans Evol Comput 2002;6(2):182–97. <https://doi.org/10.1109/4235.996017>.
- [10] Jensen NO. A note on wind generator interaction. . Riso-M-2411. Denmark: Riso National Laboratory; 1983.
- [11] Grady SA, Hussaini MY, Abdullah MM. Placement of wind turbines using genetic algorithms. Renew Energy 2005;30(2):259–70. <https://doi.org/10.1016/j.renene.2004.05.007>.
- [12] Kim T, Song J, You D. Optimization of a wind farm layout to mitigate the wind power intermittency. Appl Energy 2024;367:123383. <https://doi.org/10.1016/j.apenergy.2024.123383>.
- [13] Dincer AE, Demir A, Yilmaz K. Multi-objective turbine allocation on a wind farm site. Appl Energy 2024;355:121592. <https://doi.org/10.1016/j.apenergy.2023.121592>.
- [14] Reddy SR. Wind farm layout optimization (WindFLO): an advanced framework for fast wind farm analysis and optimization. Appl Energy 2020;269:115090. <https://doi.org/10.1016/j.apenergy.2020.115090>.
- [15] Mittal P, Mitra K. In search of flexible and robust wind farm layouts considering wind state uncertainty. J Clean Prod 2020;268:122300. <https://doi.org/10.1016/j.jclepro.2020.122300>.
- [16] Song J, Kim T, You D. Wind farm layout optimization with uncertain wind condition. Energy Convers Manag 2022;265:115783. <https://doi.org/10.1016/j.enconman.2022.115783>.
- [17] Schwarz G. Estimating the dimension of a model. Ann Stat 1978;6(2):461–4. <https://doi.org/10.1214/aos/1176344136>.
- [18] Zhang J, Zhao X. Quantification of parameter uncertainty in wind farm wake modeling. Energy 2020;196:117065. <https://doi.org/10.1016/j.energy.2020.117065>.
- [19] Pandit RK, Infield D, Kolios A. Gaussian process power curve models incorporating wind turbine operational variables. Energy Rep 2020;6:1658–69. <https://doi.org/10.1016/j.egyr.2020.06.018>.
- [20] Ouyang T, Kusiak A, He Y. Modeling wind-turbine power curve: a data partitioning and mining approach. Renew Energy 2017;102(A):1–8. <https://doi.org/10.1016/j.renene.2016.10.032>.
- [21] Gao X, Yang H, Lu L. Optimization of wind turbine layout position in a wind farm using a newly-developed two-dimensional wake model. Appl Energy 2016;174: 192–200. <https://doi.org/10.1016/j.apenergy.2016.04.098>.
- [22] Ti Z, Deng XW, Yang H. Wake modeling of wind turbines using machine learning. Appl Energy 2020;257:114003. <https://doi.org/10.1016/j.apenergy.2019.114003>.
- [23] Katic I, Hojstrup J, Jensen NO. A simple model for cluster efficiency. Proc European wind energy conference and exhibition. Rome. 1986. p. 407–10.
- [24] Bertsimas D, Sim M. The price of robustness. Oper Res 2004;52(1):35–53. <https://doi.org/10.1287/opre.1030.0065>.
- [25] Ben-Tal A, Nemirovski A. Robust solutions of uncertain linear programs. Oper Res Lett 1999;25(1):1–13. [https://doi.org/10.1016/S0167-6377\(99\)00016-4](https://doi.org/10.1016/S0167-6377(99)00016-4).
- [26] Zhai J, Wang S, Guo L, Jiang Y, Kang Z, Jones CN. Data-driven distributionally robust joint chance-constrained energy management for multi-energy microgrid. Appl Energy 2022;326:119939. <https://doi.org/10.1016/j.apenergy.2022.119939>.
- [27] Sainz E, Llombart A, Guerrero JJ. Robust filtering for the characterization of wind turbines. Energy Convers Manag 2009;50(9):2136–47. <https://doi.org/10.1016/j.enconman.2009.04.036>.
- [28] IRENA. Renewable power generation costs in 2023. Abu Dhabi: International Renewable Energy Agency; 2024. <https://www.irena.org/Publications/2024/Sept/Renewable-Power-Generation-Costs-in-2023>.
- [29] Zitzler E, Thiele L, Laumanns M, Fonseca CM, da Fonseca VG. Performance assessment of multiobjective optimizers: an analysis and review. IEEE Trans Evol Comput 2003;7(2):117–32. <https://doi.org/10.1109/TEVC.2003.810758>.
- [30] Thomas JJ, Ning A. A comparison of eight optimization methods applied to a wind farm layout optimization problem. Wind Energy Sci 2023;8:865–85. <https://doi.org/10.5194/wes-8-865-2023>.
- [31] Beyer H-G, Sendhoff B. Robust optimization — a comprehensive survey. Comput Methods Appl Mech Eng 2007;196(33–34):3190–218. <https://doi.org/10.1016/j.cma.2007.03.003>.
- [32] Faes M, Broggi M, Patelli E, Govers Y, Mottershead J, Beer M, et al. A multivariate interval approach for inverse uncertainty quantification with limited experimental

- data. Mech Syst Signal Process 2019;118:534–48. <https://doi.org/10.1016/j.ymssp.2018.08.050>.
- [33] Faes MGR, Daub M, Marelli S, Patelli E, Beer M. Engineering analysis with probability boxes: a review on computational methods. Struct Saf 2021;93:102092. <https://doi.org/10.1016/j.strusafe.2021.102092>.
- [34] van Mierlo C, Persoons A, Faes MGR, Moens D. Robust design optimisation under lack-of-knowledge uncertainty. Comput Struct 2023;275:106910. <https://doi.org/10.1016/j.compstruc.2022.106910>.



DSPACE

<https://dspace.org/>

Robust multi-objective wind farm layout optimisation under per-turbine cost uncertainty

Pandit, Ravi; Feng, Yifan

2026-12

Elsevier

<https://doi.org/10.1016/j.apenergy.2026.128437>

Pandit R, Feng Y. (2026) Robust multi-objective wind farm layout optimisation under per-turbine cost uncertainty. Applied Energy, Volume 424, December 2026, Article number 128437