



Physics-regularised neural network surrogates for fine-scale wind and pressure fields over complex terrain and structures

Jimmy Gasser¹, Yael Frischholz¹, and Michael Lehning^{1,2}

¹Laboratory of Cryospheric Sciences, Environmental Engineering Institute, École Polytechnique Fédérale de Lausanne, 1950 Sion, Switzerland

²WSL Institute for Snow and Avalanche Research SLF, 7260 Davos, Switzerland

Correspondence: Jimmy Gasser (jimmy.gasser@epfl.ch)

Abstract. Detailed wind and pressure fields around structures in complex terrain are needed for renewable energy siting or structural load estimates, but Reynolds-averaged Navier-Stokes (RANS) simulations remain too expensive for repeated siting and design studies. This work develops a neural network based surrogate of steady-state OpenFOAM simulations that predicts the three-dimensional wind velocity and kinematic pressure distribution over terrain domains of up to $3 \times 3 \text{ km}^2$ size at 30 m resolution and, in a single framework, the flow and pressure fields around solar installations and wind turbines placed in this terrain at 0.5 m resolution. Two model families, a point-wise hybrid model combining a coordinate multilayer perceptron with two-dimensional convolutional terrain encoders, and a three-dimensional grid based U-Net, are compared. Both are designed as a two-stage cascade in which a terrain-scale background model is corrected by a structure-scale residual refiner. Both are trained as physics-regularised surrogates, i.e. not only on the CFD reference data but also with penalties on the violations of the mass and momentum conservation as well as on departures from the prescribed inflow profile, outlet pressure and impermeable lateral and top boundaries. With a training budget of 292 domains, the final models reach wind speed normalised root mean square error (RMSE) values of around 0.16 for global background and of around 0.17 for regions of interest, while pressure normalised RMSE is around 0.24 for global domain and 0.12 for regions of interest. Since the kinematic pressure is physically defined only up to a constant, removing the arbitrary per-case offset with a gauge correction reduces these errors to about 0.18 – 0.20 and 0.09 respectively. Both models reach comparable values. Matched data-only comparisons show that the contribution of the physics residuals is mass- and momentum-conservation consistency rather than supervised accuracy. The wind speed error differences lie within seed-to-seed spread, while the continuity and momentum residuals are roughly halved in the structures regions and reduced by 15 to 65% at terrain scale with physics residuals. Starting from 120 simulations, the training set is expanded to 292 domains by active learning, reducing the terrain-scale test error by about 30% (wind speed normalised RMSE from 0.26 to 0.18). At a matched budget, active learning score-based selection outperforms random selection. The two model families have a trade-off, where the hybrid attains the best aggregate errors and flexible point-wise inference, whereas the U-Net better resolves the channelling jets between closely spaced structures that volume-averaged metrics barely register. The surrogate returns full fields in seconds, making repeated micro-siting, layout comparison and preliminary loading studies practical.



25 1 Introduction

The transition towards renewable energy requires wind information that is both accurate and sufficiently fast to support optimised resource assessment, site selection and engineering needs. This is particularly true in countries such as Switzerland, where many high potential renewable energy sites are located in mountainous, complex terrain (van Schaik et al., 2026). Fine, metre-scale wind predictions are also relevant beyond energy plant siting, for example for the estimation of structural loads or snow transport (Mott and Lehning, 2010; Lehning et al., 2008), the latter affecting PV yield through snow cover and albedo (Kahl et al., 2019).

Global reanalysis and numerical weather prediction (NWP) models describe synoptic and mesoscale atmospheric state, but their kilometre-scale grids cannot explicitly resolve important terrain features. ERA5 (Hersbach et al., 2020) for example provides a consistent global atmospheric reconstruction at a 31 km scale, while NWPs such as COSMO (Doms and Baldauf, 2018) or ICON (Zängl et al., 2015) resolve regional atmospheric dynamics at finer, yet still kilometre resolution. Even the most recent high-resolution regional frameworks of these models remain too coarse to capture effects of small scale topography. In complex terrain such as mountainous regions, the local flow can indeed be dominated by terrain-induced effects like valley channelling, sheltering or ridge acceleration (Raderschall et al., 2008; Berg et al., 2011). An evaluation of COSMO-1 over the Swiss Alps has shown that these limitations can lead to errors at mountain stations (Krüy et al., 2018).

Several downscaling approaches have therefore been developed to reconstruct terrain-scale wind fields from coarser atmospheric information. For example, HICAR (Reynolds et al., 2023) is an intermediate-complexity atmospheric model designed to dynamically downscale kilometre-scale atmospheric forcing to the hectometre-scale. By simplifying parts of the atmospheric dynamics while retaining relevant parameterisations, it can represent terrain-induced effects at a substantially lower cost than NWP models like WRF (Skamarock et al., 2019). Other approaches to the downscaling of numerical wind fields make use of statistical and machine learning models. An example is Wind-Topo (Dujardin and Lehning, 2022), which reconstructs near-surface wind fields at 50 m resolution directly from high-resolution topography and coarse COSMO-1 atmospheric states as well as station observations in Switzerland. For renewable energy planning, predictions must be accurate and cheap enough to evaluate large regions and multiple candidate locations. Wind atlases, also based on downscaling approaches, have therefore been developed as efficient tools, with for example the European Wind Atlas (Troen and Petersen, 1989), the Global Wind Atlas (Davis et al., 2023), or the Swiss Wind Atlas (Swiss Federal Office of Energy (SFOE), 2023). The latter provides nationwide wind speed on a 100 m grid. These products allow fast region identification, but they struggle to represent separated flow, sharp pressure gradients or recirculation. They moreover only give average wind climates and do not resolve seasonal or sub-seasonal variations, which is required for energy system modelling (Lehning et al., 2025).

The need for accurate terrain-scale wind predictions is particularly true in Switzerland, where the Energy Strategy 2050 sets targets for a strong expansion of renewable energy generation (Swiss Federal Office of Energy (SFOE), 2024), and where many promising sites are located in complex alpine terrain. High alpine photovoltaic installations can provide enhanced winter production, when electricity demand is higher (Kahl et al., 2019). Wind power in complex terrain may further complement this seasonal production pattern (Krüy et al., 2017; Dujardin et al., 2017, 2021; Lehning et al., 2025). However, assessing



alpine wind resources in such regions remains difficult and often needs validation through measurements. Recent studies have combined LiDAR campaigns and machine learning to estimate wind resources in the Swiss Alps (Kristianti et al., 2023, 2024) and LiDAR measurements with computational fluid dynamics (CFD) to reconstruct three dimensional flow over a mountain pass (van Schaik et al., 2026). These approaches provide valuable information at particular sites, but they depend on local measurement campaigns, and their generalisation to other sites, let alone to many candidate sites at once, remains to be shown.

The downscaling models and measurement based studies discussed above address the influence of terrain on the atmospheric flow. However, renewable energy installations introduce an additional, finer spatial scale. Solar panels and wind turbines in particular generate stagnation regions, wakes and recirculation occurring over distances of metres, which cannot be recovered by interpolating a terrain-scale wind field. Resolving these metre-scale effects is important not only for structural loading estimates and installation siting, but also for the modelling of processes such as snow transport. Indeed, local wind fields influence snow deposition patterns (Mott and Lehning, 2010), which can block or expose parts of an installation and modify photovoltaic production through changes in snow cover and surface albedo (Kahl et al., 2019). Fine scale wind prediction links the atmospheric resource assessment problem to a local engineering one that is relevant to renewable energy performance.

Computational fluid dynamics (CFD) provides a physically based framework for connecting these two scales. By numerically solving approximations of the Navier-Stokes equations, CFD can represent atmospheric flow in complex topography while simultaneously accounting for explicit structure geometries. Reynolds-averaged Navier-Stokes (RANS) simulations are commonly used when time-averaged wind and pressure fields are required at a computational cost below that of scale-resolving turbulence methods such as large eddy simulations (LES). OpenFOAM (Weller et al., 1998) provides an established open-source framework for constructing such simulations and has been widely applied to atmospheric and engineering flows. Although CFD can represent both terrain and structure induced effects, its computational cost remains a limitation. Additionally, each simulation requires geometry preparation, mesh generation, numerical solving and post-processing. This becomes particularly demanding for mountainous regions containing small structures, as the mesh must extend over kilometres while retaining metre-scale resolution around the structure of interest. Repeating this workflow for various terrains, structure geometries, wind directions, structure layouts and inflow conditions is therefore too expensive for rapid engineering assessment.

Neural network surrogates offer a possible way to retain the spatial details provided by the CFD while greatly reducing the cost of prediction. Instead of solving the governing equations from the start for every case, a surrogate learns the relationship between the domain, the atmospheric conditions and the resulting flow fields from a database of previously computed CFD solutions. Once trained, such models can predict new cases within seconds, making them potentially suitable for structure layout comparisons or repeated assessments. Early work demonstrated that neural networks could approximate steady CFD solutions from geometrical descriptions (Guo et al., 2016), and encoder-decoder architectures such as U-Net (Ronneberger et al., 2015) have been widely used for fluid flow surrogates (Thuerey et al., 2020). Surrogate models have also been developed for flows involving explicit built geometries (Calzolari and Liu, 2021). Aly and Clarke (2023) trained a machine-learning surrogate of RANS simulations for the wind-resistant design of ground-mounted solar panels, predicting the velocity and pressure distributions around a single panel with a reported speed-up of four orders of magnitude. The approach however considers an isolated panel over simplified ground, rather than structures embedded in kilometre-scale terrain. A recent graph-



based generative approach reconstructs urban wind fields on irregular meshes from sparse sensor measurements (Giral et al., 2026). The study demonstrates that a machine-learning model can represent structure induced flow effects, although it focuses on limited local domains rather than combining them with kilometre-scale terrain flow. At the terrain-scale, FuXi-CFD (Lin et al., 2026) provides a recent example of a CFD-trained atmospheric surrogate. The Vision Transformer based neural network model learns terrain informed 3D wind reconstruction from CFD-generated training data at 30 m resolution. It demonstrates that learned surrogates can efficiently recover terrain induced flow variations over kilometre-scale domains. However, the model's resolution remains too coarse to resolve structure induced flows at the metre-scale. Another model, WindSeer (Achermann et al., 2024), predicts volumetric three-dimensional wind over complex terrain from only a few sparse, noisy measurements, using a CFD trained neural network. While reducing inference time, the cost of production of the training dataset remains the main limitation to these data-intensive architectures. For example, FuXi-CFD was trained using more than 300 000 CPU hours for over 12 000 CFD simulations, while WindSeer used about 10 000 CPU hours for more than 7 000 simulations. Although these products address different tasks, resolutions and outputs, so that their training budgets are not directly comparable to the present setting, they show that the database generation is the dominant cost of building a surrogate. One way to reduce this cost is to adapt the training strategy and choose the simulations to run more carefully. With active learning (AL), Rygiel et al. (2025) for instance query new CFD samples using geometric diversity, ensemble uncertainty and physics-residual criteria, reducing the number of simulations needed for a given accuracy by up to 50%. Adding physical information to the training objective is another strategy to reduce the degree of reliance on data. Physics-informed neural networks (PINNs) (Raissi et al., 2019) train a network to satisfy the governing partial differential equations by minimising their residuals at collocation points, with sparse or no reference data. When a database of reference solutions is available, the same residuals can instead be added to the supervised error of a data-driven surrogate, as done in physics-informed neural operators (Li et al., 2024; Wang et al., 2021). The network is then penalised not only when its predictions differ from the reference CFD fields, but also when they violate physical relations such as mass and momentum conservation. Physics-informed formulations have previously been applied to the RANS equations (Eivazi et al., 2022) and to surrogate modelling of fluid flows. In the present application, CFD reference fields can be generated, but producing a database covering every combination of terrain containing structures with different orientations is computationally infeasible. The governing equation residuals can therefore be interpreted as an additional physically motivated regulariser. They constrain the predicted flow between supervised samples and may consequently allow a model to reach a given predictive accuracy using fewer CFD simulations than a purely data-driven surrogate. We refer to this second approach as physics-regularised.

The approaches discussed here form two parallel multiscale modelling chains. On the numerical side, atmospheric information progresses from global reanalysis, such as ERA5 (Hersbach et al., 2020), through regional and intermediate-complexity atmospheric models such as HICAR (Reynolds et al., 2023) to metre-scale CFD with OpenFOAM (Weller et al., 1998), spanning roughly 10^4 to 10^0 m. On the machine-learning side, a corresponding hierarchy is emerging. Global data driven forecasting models such as FourCastNet (Pathak et al., 2022) are followed by terrain-scale downscaling models such as Wind-Topo or FuXi-CFD. At the smallest end of this hierarchy, machine-learning models have been developed for flow around buildings for example, but these generally assume prescribed inflow conditions and consider limited local domains. A corresponding surro-



gate that connects kilometre-scale terrain-induced flow with metre-scale disturbances around structures remains comparatively
130 underdeveloped. The model presented in this work is intended to address this final link in the hierarchy.

This work presents a surrogate that resolves the terrain-scale wind field over domains of up to $3 \times 3 \text{ km}^2$ and, within the same
framework, the metre-scale flow corrections around isolated solar structures, wind turbines, and regular arrays of structures.
We thus present a single model that combines two regimes within one framework, with 30 m resolution flow over kilometres,
and 0.5 m resolution structure-induced flows.

135 The study addresses four main questions structuring the results:

- How accurately and how fast can a single CFD surrogate model predict both the terrain-scale flow and the metre-scale
flow around the structures?
- What do the physics residual losses contribute in this setting?
- How can such a surrogate be designed in terms of architecture, loss formulation and training design?
- 140 – Can active learning reduce the simulation budget?

2 Methodology

2.1 CFD dataset

The reference data are steady-state RANS simulations produced with OpenFOAM (simpleFoam, $k-\varepsilon$ closure with atmospheric
rough-wall functions, kinematic viscosity $\nu = 1.5 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$) (Weller et al., 1998). Each simulation provides the velocity
145 field ($\mathbf{u}$), the kinematic pressure (p), the turbulent kinetic energy (k), its dissipation rate (ε) and the turbulent kinematic viscosity
(ν_t). Each case is generated automatically with the following steps. A square domain with a side length between 0.5 and 3 km is
extracted from the swissALTI3D digital elevation model at 2 m resolution at a chosen Swiss location and oriented such that the
inflow is aligned with the $+x$ axis. The domain is then meshed with snappyHexMesh with progressive refinement towards
the ground and towards structure surfaces (background cells of 5-30 m, refined to metre-scale at the ground and to decimetre-
150 scale on structures, with meshes up to 3×10^7 cells for domains with multiple single structures). The inlet prescribes a neutral
logarithmic atmospheric boundary layer profile with reference speed $U_{\text{ref}} \in [5, 15] \text{ m s}^{-1}$ at $Z_{\text{ref}} = 20 \text{ m}$ and aerodynamic
roughness length $z_0 = 0.1 \text{ m}$, while the outlet fixes the kinematic pressure reference $p = 0$. The lateral and top boundaries are
treated as slip surfaces whereas the terrain and structure surfaces are modelled as no-slip walls. Simulations are run for at most
5000 SIMPLE iterations and stopped earlier if the normalised residuals of all solved variables, i.e. velocity, pressure, k and ε ,
155 fall below 10^{-4} . Some domains (9%) reach the iteration limit with residuals that have plateaued at low values. This is common
for steady RANS simulations of complex terrain and separated flows around structures, where small, inherently unsteady
regions prevent a further decrease of the residuals. Following established guidelines (Franke et al., 2007; Blocken, 2015), such
solutions are retained if the residuals have levelled off and the flow fields no longer oscillate over the final iterations.



The sensitivity of the solution to the mesh was assessed on a representative, $3 \times 3 \text{ km}^2$ complex terrain domain with structures. The background cell size was varied by a factor of two, the cells of the refinement regions being defined relatively to these background cells and therefore scaled by the same factor. Three meshes were tested, with background cells of 60 m (1.1×10^6 total cells), 30 m (5.5×10^6 total cells) and 15 m (23×10^6 total cells). A fourth mesh with 7.5 m background cells (150×10^6 total cells) was rejected, as such domains are not affordable for a large dataset creation. The volume-weighted root mean square (RMS) difference in wind speed between successive meshes, normalised by the reference wind speed of the domain, decreases from 0.15 to 0.003 over the whole domain, and from 0.27 to 0.011 in the lowest 10 m above the terrain. The coarsest mesh systematically under-estimates the near-ground wind speed, which is no longer the case between the two finer meshes. The mesh with 30 m background cells, corresponding to 100 background cells per domain side, was thus retained for the dataset creation. The same number of background cells is used for smaller domains.

Structures are placed in the terrain as solid bodies from a library of eight photovoltaic structure types (flat or inclined table-mounted panels, "Helioplant®" units and concentrators, with characteristic heights of 2-6 m) and two wind-turbine geometries (20 m and 80 m total height), the latter represented with a stationary rotor. Single-structure cases contain one to five structures that are mutually isolated in the domain while multi-structure cases contain regular grids of up to 10×10 identical structures with varying spacing, row orientation and panel yaw, in which the flow around each structure interacts with its neighbours.

Each simulation is resampled onto two regular grids used for training. A global grid covers the full domain at 30 m horizontal spacing with terrain-anchored vertical bands refined to 2 m near the surface. At this resolution of 30 m, comparable to models such as FuXi-CFD, the structures are not explicitly resolved. When structures are present, one or more regions of interest (ROIs) grids centred on the structures are additionally exported at 0.5 m resolution, extending several structure heights upstream, downstream and above. For isolated single-structure cases, an ROI is defined for each structure, whereas for multi-structure cases, one single ROI is defined for the whole array of structures.

The original dataset consists of 120 hand-designed domains, divided into three families, with 40 complex terrain-only cases covering basic configurations such as along- and cross-valley flow, up- and downhill flow, ridges, hills and peaks, 40 single-structure cases and 40 multi-structure cases or arrays. Note that structures are, themselves, located in real, complex terrain domains. The 120 cases are split into 84 training, 18 validation and 18 test domains, with six domains of each family in the validation and in the test set. Both the validation and the test sets are kept frozen during the entire model development, while the training set is then gradually expanded by active learning (Section 3.5). In active learning, trained models are used for predictions on a large set of unsimulated candidate domains. A subset of domains with the worst prediction (ranked by the selection score of section 3.5, as no CFD data exists for these candidates) is then selected, meshed, simulated and added to the training data. The methodology follows work presented by Rygiel et al. (2025). Different active learning iterations have brought the dataset from 120 to a final dataset of 292 domains (256 training, 18 validation, 18 test; 108 complex terrain only, 91 single-structure, 93 multi-structures). Representative examples of obtained CFD result for each family (complex terrain only, single-structure, multi-structure) are shown in figure 1.

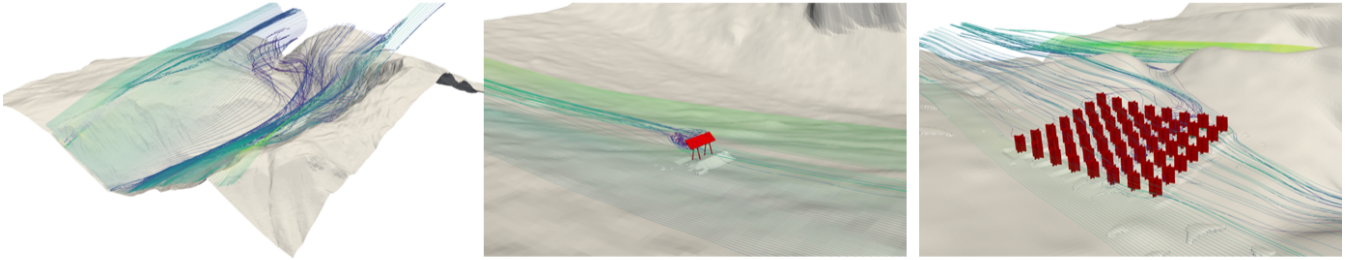


Figure 1. Obtained CFD data for a complex terrain only case (left), a single-structure case, here an inclined solar panel (middle) and a multi-structure case, here an array of Helioplant® units (right). The structures are highlighted in red, while local wind streamlines are shown in arbitrary colour scales. Terrain data: swissALTI3D © swisstopo.

2.2 Model architecture families

Two neural-network model families have been developed, allowing comparison of prediction accuracy but also of inference flexibility on the same learning problem. The first family is a hybrid point-wise model that can be queried at arbitrary coordinates, whereas the second is a grid-based U-Net that predicts spatially coupled voxel volumes. Both share the same input data and predict the same output fields, but differ in how they represent and evaluate the three-dimensional flow. They were chosen because they represent two established ways of building a CFD surrogate: coordinate-based networks, which evaluate the flow at arbitrary points and are the natural host for differential-equation residuals (Raissi et al., 2019) and grid-based convolutional encoder-decoders, which are widely used for steady flow surrogates (Guo et al., 2016; Thuerey et al., 2020). Attention-based architectures were not considered at this stage.

The model inputs consist of local features and two dimensional conditioning maps. The local features describe either the point queried by the hybrid model or the voxel processed by the U-Net, depending on the model family. These local features contain the coordinates (x, y, z) , the height above terrain $z_{\text{rel}} = z - z_{\text{terrain}}$, a smooth near-ground transform of z_{rel} , the local terrain elevation and slope, terrain aspect represented by its sine and cosine, the inflow conditions $(U_{\text{ref}}, Z_{\text{ref}}, z_0)$, the inflow direction unit vector and the three domain dimensions. The roughness length is supplied as $\log_{10}(z_0)$. All quantities are expressed in the canonical frame introduced in Section 2.1, in which the domain is oriented so that the inflow is aligned with the $+x$ axis. Wind direction is therefore absorbed into the domain orientation and is not treated as a varying model input. Within structure regions of interest, an additional local feature is the signed distance to the nearest structure surface. The global and ROI inputs contain 18 and 19 local features, respectively.

The two dimensional conditioning maps provide spatial context to the model that cannot be inferred from an isolated query point. A four-channel terrain map contains centred elevation, slope, and aspect represented by its sine and cosine. In the structure-scale refiner, an additional context map describes the surrounding structure geometry using footprint and height fields, signed distance to structure walls, streamwise and cross-stream offsets relative to the inflow and local structure density. These quantities are computed entirely from the structure geometry and inflow orientation and do not use the CFD solution.



215 Local features and conditioning maps are complementary, as quantities such as elevation are provided both as values at the prediction location and as spatial fields describing their surroundings.

Both model families return the three velocity components (u_x, u_y, u_z) and the kinematic pressure p at each output location. They are moreover organised as the same two-stage cascade. In Stage 1, the model is trained only on the global fields at 30 m horizontal resolution and provides a terrain-scale background prediction. Note that domains containing structures contribute to their global fields as well, and the influence of the structures on the terrain-scale flow is therefore already present in the Stage 1 training data. Once trained, the model weights are frozen. Stage 2 is a structure-scale refiner trained on the 0.5 m ROI fields. It receives the ROI features, the structure-context information and the Stage 1 prediction interpolated to the ROI locations, and predicts a residual correction that is added to the background field. The residual output layer is initialised to zero, so that Stage 2 training begins from the Stage 1 prediction. An edge penalty additionally constrains the residual towards zero within a band along the ROI boundaries, allowing the refined field to merge smoothly with the surrounding terrain-scale prediction.

The first model family combines a coordinate multilayer perceptron (MLP) with two dimensional convolutional encoders, and is therefore named hybrid in the rest of the paper. In Stage 1, the two dimensional terrain map is processed by a dilated 2D CNN with width 32, depth 5, and dilation factors increasing up to 32, giving a receptive field that covers the full terrain domain. The encoded terrain features are bilinearly interpolated at the horizontal location (x, y) of each query point. They are concatenated with the local input features and a Fourier encoding of the three-dimensional coordinates before being passed to a residual MLP head of width 256 and depth 4, which predicts (u_x, u_y, u_z, p) . Fourier features map the three coordinates (x, y, z) through sines and cosines of F random frequency vectors, drawn once at initialisation from a Gaussian distribution of standard deviation σ (the bandwidth), which lets a coordinate network represent higher frequency spatial variation rather than only smooth functions (Tancik et al., 2020). Four frequencies with bandwidth $\sigma = 1$ are used in Stage 1, and eight with bandwidth $\sigma = 2$ in the structure-scale refiner. Because points are evaluated independently, the hybrid model can be queried at arbitrary coordinates and resolutions, with peak memory controlled through the query batch size rather than by the dimensions of a complete output volume. The Stage 1 hybrid model contains approximately 1.2 million trainable parameters. Its Stage 2 refiner uses the dedicated ROI terrain and structure encoders together with a point-wise residual head and contains approximately 0.6 million parameters.

240 The second family is a three-dimensional adaptation of the U-Net architecture (Ronneberger et al., 2015), called U-Net in the rest of the study. The Stage 1 network has four resolution levels, with channel widths increasing from 32 to 256, and maps a voxel tensor of all local input features to the four output fields on the structured grid. A dilated 2D terrain encoder with the same width, depth and dilation pattern as in the hybrid family provides full-domain horizontal context. Its encoded features are broadcast along the vertical axis and concatenated with the three-dimensional voxel features. The U-Net predicts complete volumes using spatially coupled convolutional kernels and is consequently tied to the prescribed voxel resolution. The Stage 1 U-Net contains approximately 11 million trainable parameters. Its Stage 2 refiner is a second 3D U-Net operating on the ROI grids and contains approximately 5.6 million parameters.

The two families therefore share the same inputs, outputs and terrain-to-structure cascade. Their principal difference lies in the representation of the flow field, as the hybrid model treats output locations as independent queries conditioned on encoded



250 spatial context, whereas the U-Net directly couples neighbouring voxels through three-dimensional convolutions. Table 1 summarises the differences in inference behaviour. Figures B2 and B3 in Appendix B give a visual representation of both models.

Table 1. Comparison of the two model architecture families. Both families use the same physical inputs and output variables and follow the same two-stage terrain-to-structure cascade.

	Hybrid point-wise model	Grid U-Net model
Prediction unit	Independent query points	Complete voxel volumes
Main backbone	Coordinate MLP	Three dimensional U-Net
Spatial context	Sampled 2D encoded maps	2D context with 3D convolutions
Coupling between outputs	No point to point coupling	Coupled by convolutions
Evaluation locations	Arbitrary coordinates	Prescribed structured grid
Output resolution	Selectable at inference	Fixed by the voxel spacing
Parameters, Stage 1	≈ 1.2 M	≈ 11 M
Parameters, Stage 2	≈ 0.6 M	≈ 5.6 M

2.3 Model losses

Both the hybrid and the U-Net models use the same training losses, which combine data-driven losses, physics residual losses
 255 and boundary condition losses. We denote $\hat{\mathbf{y}} = (\hat{u}_x, \hat{u}_y, \hat{u}_z, \hat{p})$ the predicted fields, $\mathbf{y}$ the corresponding CFD reference. Both are expressed in normalised output space when evaluating the supervised loss. For the physics and boundary-condition losses, the predicted fields are transformed back to physical units. We denote by $\mathbf{u} = (u_x, u_y, u_z)$ the velocity and by p the kinematic pressure in physical units.

The supervised loss uses a Charbonnier penalty (Charbonnier et al., 1994) applied channel-wise and averaged over the N
 260 sampled points,

$$\mathcal{L}_{\text{data}} = \frac{1}{N} \sum_{i=1}^N w_i \sum_c \lambda_c \rho(\hat{y}_{i,c} - y_{i,c}), \quad \rho(e) = \sqrt{e^2 + \varepsilon^2} - \varepsilon, \quad (1)$$

where c indexes the four output channels, λ_c are channel weights (for example for giving pressure a dedicated weight p_w), ε is a small constant ($\varepsilon = 10^{-3}$ here), and w_i is an optional per-point weight that can emphasise near-wall, wake or low-speed regions in the structure refiner. The Charbonnier penalty behaves as an L_2 loss for small residuals $|e| \ll \varepsilon$, where $\rho(e) \approx e^2/2\varepsilon$, and as
 265 an L_1 loss for large residuals $|e| \gg \varepsilon$, where $\rho(e) \approx |e|$. With $\varepsilon = 10^{-3}$ and normalised targets, nearly all residuals exceed ε , so the penalty acts in practice as a smooth L_1 loss, which limits the influence of the largest errors on the training. This aims to limit the known tendency of pure L_2 losses to oversmooth sharp gradients, while remaining robust to large local error.



Two physics losses penalise departures from the steady, incompressible Reynolds-averaged Navier-Stokes (RANS) equations. The continuity loss penalises departures from mass conservation:

$$\mathcal{L}_{\text{div}} = \frac{1}{M} \sum_{j=1}^M \left(\frac{\nabla \cdot \mathbf{u}_j}{s_{\text{div}}} \right)^2, \quad \nabla \cdot \mathbf{u} = \partial_x u_x + \partial_y u_y + \partial_z u_z, \quad (2)$$

The momentum loss penalises departures from the steady RANS momentum balance:

$$\mathcal{L}_{\text{mom}} = \frac{1}{M} \sum_{j=1}^M \frac{\|\mathbf{r}_j\|^2}{s_{\text{mom}}^2}, \quad \mathbf{r} = (\mathbf{u} \cdot \nabla) \mathbf{u} + \nabla p - \nabla \cdot [\nu_{\text{eff}} (\nabla \mathbf{u} + \nabla \mathbf{u}^\top - \frac{2}{3} (\nabla \cdot \mathbf{u}) \mathbf{I})]. \quad (3)$$

The two residuals are evaluated by finite differences on random structured patches of the export grids and averaged over the M fluid cells they contain. Evaluating the residuals by finite differences on structured samples rather than by automatic differentiation at isolated collocation points follows previous finite difference-based physics-informed formulations (Jiang et al., 2023). It avoids differentiating twice through the network graph, which proved slower and less stable in early experiments, and comes at no additional cost since the reference data are exported on structured grids. The scales $s_{\text{div}} = U_{\text{ref}}/Z_{\text{ref}}$ and $s_{\text{mom}} = U_{\text{ref}}^2/Z_{\text{ref}}$ are the velocity-gradient and advective-acceleration scales of the inflow that normalise the two residuals to comparable, dimensionless magnitudes. The momentum loss uses the effective viscosity $\nu_{\text{eff}} = \nu + \nu_t$, using the kinematic viscosity $\nu = 1.5 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$ with the turbulent kinematic viscosity ν_t provided by the CFD reference data. The momentum operator therefore remains consistent with the turbulence closure used to generate the reference CFD fields and acts as a physically grounded regulariser that penalises departures from the modelled momentum balance. The turbulent viscosity entering the residual is taken from the reference simulation. It is therefore available for every training and test domain, all of which are simulated, but not for a domain that has never been simulated, as the surrogate itself does not predict ν_t or any turbulence related variable yet. This limitation is discussed in section 3.6.

Additional terms penalise departures from the domain boundary conditions of the global grid for Stage 1, with the prescribed logarithmic inflow profile at the inlet, the reference kinematic pressure at the outlet and the impermeability of the lateral and top slip boundaries:

$$\mathcal{L}_{\text{bc}} = w_{\text{in}} \mathcal{L}_{\text{inlet}} + w_{\text{out}} \mathcal{L}_{\text{outlet}} + w_{\text{side}} \mathcal{L}_{\text{side}} + w_{\text{top}} \mathcal{L}_{\text{top}}, \quad (4)$$

with $(w_{\text{in}}, w_{\text{out}}, w_{\text{side}}, w_{\text{top}})$ fixed relative weights. The four loss terms are written out in Appendix C.

The two cascade stages are trained sequentially, each with its own weighted objective. Stage 1 is trained on the global 30 m field domain (superscript g) and combines the supervised, physics and boundary terms:

$$\mathcal{L}^{(1)} = \lambda_{\text{data}}^g \mathcal{L}_{\text{data}}^g + \lambda_{\text{div}} \mathcal{L}_{\text{div}}^g + \lambda_{\text{mom}} \mathcal{L}_{\text{mom}}^g + \lambda_{\text{bc}} \mathcal{L}_{\text{bc}}. \quad (5)$$

After Stage 1 is trained and frozen, Stage 2 is trained on the 0.5 m ROI fields (superscript r) using the supervised and physics terms together with the edge penalty $\mathcal{L}_{\text{edge}}$ that anchors the residual correction to zero along the ROI borders:

$$\mathcal{L}^{(2)} = \lambda_{\text{data}}^r \mathcal{L}_{\text{data}}^r + \lambda_{\text{div}} \mathcal{L}_{\text{div}}^r + \lambda_{\text{mom}} \mathcal{L}_{\text{mom}}^r + \lambda_{\text{edge}} \mathcal{L}_{\text{edge}}. \quad (6)$$



To isolate the contribution of the physics residuals, each stage is also trained as a matched data-only counterpart with the differential-equation weights set to zero,

$$\mathcal{L}_{\text{nophys}}^{(s)} = \mathcal{L}^{(s)} \Big|_{\lambda_{\text{div}}=\lambda_{\text{mom}}=0}, \quad s \in \{1, 2\}. \quad (7)$$

300 2.4 Training and model selection

Each stage of both models is trained with the AdamW optimiser (Loshchilov and Hutter, 2019), a one-cycle cosine learning-rate schedule (Smith and Topin, 2019) and up to 500 epochs with early stopping. The hybrid model samples, per training case and epoch, a batch of random global points together with the structured physics patches and boundary points required by the loss. The U-Net model is instead trained on random $32 \times 32 \times 16$ patches of the global (Stage 1) and ROI (Stage 2) grids, with
 305 patch origins biased towards the ground and the structures. Per case and epoch, the hybrid model draws 4096 points in Stage 1, about 1% of the fluid cells of a median global grid, and 16384 points per ROI in Stage 2, about 3% of a median ROI, whereas the four U-Net patches cover about 10% of a median global grid and of a median ROI grid. In both families the Stage-2 refiner draws a targeted ROI sample weighted towards the near-wall and wake regions. Since the points and patches are redrawn at every epoch, each domain is sampled several times over during the 500 epochs, and the development experiments presented
 310 in Appendix A showed no systematic gain from denser sampling. Inputs and outputs are normalised using statistics from the training cases only.

Both models are trained with a streamwise-mirror data augmentation. As the steady RANS system is invariant under the reflection $y \rightarrow -y$ (no Coriolis force is modelled), each training case also admits a physically valid mirror image obtained by this reflection together with the sign flip $u_y \rightarrow -u_y$ (the terrain, structures and all other fields being reflected accordingly).
 315 Adding these mirrored domains doubles the effective number of training domains without additional CFD cost. The validation and test domains are not mirrored.

After each epoch, every validation global domain and ROI is predicted in full and a scalar selection score combines the wind speed and pressure normalised root mean square errors (RMSE) averaged over the relevant global and ROI scopes:

$$S = \overline{\text{nRMSE}_U} + w_p \overline{\text{nRMSE}_p}, \quad w_p = 0.3, \quad (8)$$

320 where, for one domain or ROI with N fluid cells,

$$\text{nRMSE}_U = \frac{1}{U_{\text{ref}}} \sqrt{\frac{1}{N} \sum_{i=1}^N (\|\mathbf{u}_i\| - \|\mathbf{u}_i^{\text{CFD}}\|)^2}, \quad \text{nRMSE}_p = \frac{1}{\frac{1}{2}U_{\text{ref}}^2} \sqrt{\frac{1}{N} \sum_{i=1}^N (p_i - p_i^{\text{CFD}})^2}, \quad (9)$$

and the overbar denotes the mean over the domains or ROIs of a split, each with equal weight. The same two metrics are used for all test results reported in section 3. The checkpoint with the lowest validation score S is retained. While the physics residuals contribute to model training, they do not enter the checkpoint selection score, which is based on the held-out validation error
 325 only.

The architectures, input features, loss formulation and training settings described above correspond to the final selected configuration of each model family. They result from a series of development experiments reported in Appendix A. The loss



terms were assembled incrementally at the initial 120-domain budget, starting from a purely data-driven mean squared error (MSE) model and adding the continuity residual, the momentum residual and finally the weighted Charbonnier loss. These changes brought the model error down, with the ROI pressure normalised RMSE decreasing from 0.434 to 0.130 for example. A subsequent screening of about 40 architecture, sampling and loss configurations was also performed and showed the accuracy to be only weakly sensitive to most architectural hyperparameters. Development effort was therefore concentrated on the loss formulation and on the training data, keeping the architectures close to their baselines. The contribution of the physics terms at the full training budget is assessed separately in section 3.3.

3 Results and discussion

3.1 Final model performance

We first evaluate the two final models, the hybrid and the U-Net, trained on the complete 292-domain dataset with the weighted Charbonnier loss, physics losses and streamwise-mirror augmentation, as described in section 2. Both are assessed on the frozen 18-domain test set using primarily the normalised RMSE of the wind speed magnitude ($U \equiv \|\mathbf{u}\|$) and of the kinematic pressure p , the two quantities driving validation selection.

Table 2 reports these errors for the complete two-stage models, on the Stage 1 global 30 m grid and on the Stage 2 0.5 m ROIs. The two families reach a comparable accuracy. The global wind speed normalised RMSE is 0.16 for the hybrid and 0.18 for the U-Net, i.e. 16 to 18% of each case's reference wind speed.. The global pressure normalised RMSE is higher, with 0.25 and 0.23 for the hybrid and U-Net model respectively. We also report a gauge-corrected pressure, whose value is obtained by subtracting the optimal constant offset from each test case before computing the error. This diagnostic is useful because incompressible kinematic pressure is defined only up to an additive constant, so a case-wise pressure offset can increase the raw pressure normalised RMSE without changing the spatial pressure gradients. This gauge corrected pressure normalised RMSE reaches lower values of 0.20 and 0.19 for the hybrid and the U-Net model. In the structure ROIs both families reach a wind speed normalised RMSE of 0.17 to 0.18, a pressure normalised RMSE of 0.12 to 0.13 and a gauge corrected pressure normalised RMSE of 0.09. The gauge corrected pressure can be taken as the more physically meaningful measure of the predicted pressure field, and the final model pressure error is best read from it. However for simplicity, only the raw pressure normalised RMSE will be reported in the remainder of the study.

It can be noted that both final models use the streamwise-mirror augmentation described in section 2. Relative to the matched models trained on the same 292 domains without it, the mirror augmentation lowers the wind speed test errors by up to seventeen percent depending on model and scale (hybrid ROI wind speed decreases from 0.20 to 0.17 and global wind speed decreases from 0.17 to 0.16). The ROI pressure, also the metric with the largest seed-to-seed spread, responds differently for the two model families (hybrid 0.11 to 0.12 and U-Net from 0.24 to 0.13).

Furthermore, training is stochastic through initialisation and batch sampling. To indicate the uncertainty in this performance assessment, both stages of both model families were retrained with a second seed. The wind speed, the pressure and the gauge-corrected pressure errors vary by up to approximately 0.01 between seeds at both scales. These values are ranges over two



seeds and may understate the true spread; they are quoted so that differences of comparable size are not over-interpreted in the results discussed later.

Table 2. Final two-stage models on the frozen test set. Test normalised RMSE of the wind speed and kinematic pressure (raw and gauge-corrected) on the Stage 1 global 30 m resolution grid and the Stage 2 0.5 m resolution ROIs.

Model	global U	global p	global p (gauge)	ROI U	ROI p	ROI p (gauge)
Hybrid	0.157	0.247	0.202	0.169	0.122	0.089
U-Net	0.177	0.233	0.185	0.176	0.130	0.086

Beyond the aggregate normalised RMSE, table 3 reports the signed and absolute bias of the wind speed and pressure, normalised by the same per-case reference scales as the normalised RMSE, together with the mean horizontal wind-direction error. In the bulk of the global domain both models are nearly unbiased, as the signed wind speed bias is within 1% of the reference speed. The larger mean absolute bias (5-11%) thus reflects per-case scatter that mostly cancels rather than a systematic offset. The mean direction error in the bulk is low, below 6°. The principal exception is the near-ground region, where both models systematically over-predict the wind speed (signed bias +0.09 for the hybrid, +0.18 for the U-Net) and the direction error rises to 14-19°. This over-prediction is concentrated in the first cells above terrain, where the wind speed drops most sharply towards zero at the surface. Both models tend to smooth out this sharp drop, which raises the predicted speed in the lowest cells. This bias persists even with dense data sampling near the ground. Furthermore, physics residuals offer little correction here. Indeed, at the ground boundary, numerical stencils lack underlying cells and OpenFOAM's reference momentum balance is closed by algebraic wall-function shear rather than the continuous differential operators the neural networks are trying to enforce. The contribution of the near-ground cells is moreover diluted in the volume-averaged normalised RMSE. The over-prediction could be addressed by over-weighting near-ground cells in the training objective, at the cost of de-emphasising the bulk flow. Additionally, the reference itself is least reliable in this region since the first cells above the terrain are governed by the wall-function closure where RANS is most uncertain. The near-ground error therefore combines surrogate error with the largest CFD uncertainty. In the ROIs the biases are again small, with a residual near-ground over-speed and a larger direction error for the U-Net.



Table 3. Normalised bias and direction error of the two final models on the frozen test set. "signed" is the mean signed bias (positive: over-prediction), " $|\text{bias}|$ " the mean absolute bias, both normalised by the per-case reference wind speed or pressure as for the normalised RMSE. The direction error is the mean angular error of the horizontal wind. "Near-ground" is the subset of cells closest to the terrain.

Model	Region	signed U	$ \text{bias} U$	signed p	$ \text{bias} p$	dir. err. [°]
Hybrid	Global (bulk)	+0.008	0.050	−0.023	0.116	4.6
	Global (near-ground)	+0.094	0.112	−0.014	0.136	14.3
	ROI	+0.018	0.106	−0.002	0.071	5.6
U-Net	Global (bulk)	−0.011	0.046	+0.016	0.125	5.8
	Global (near-ground)	+0.179	0.182	−0.008	0.152	19.4
	ROI	−0.049	0.087	+0.054	0.085	13.7

To complement the aggregate errors, figure 2 shows the predicted versus CFD reference density scatter plot of all output points sampled over the 18 test domains for the hybrid model. Each panel reports the coefficient of determination R^2 , the mean absolute error (MAE) and the mean bias error (MBE). The latter two are given in physical units and computed over all points pooled together, so that the largest domains and those with the highest wind speeds dominate. Reported metrics thus differ from the values of tables 2 and 3, where the metrics of each domain are first normalised by their reference scale and then averaged over the 18 domains with equal weights. For example, the negative global wind speed MBE of -0.40 m s^{-1} comes mostly from a domain with a reference wind speed of 15 m s^{-1} where the wind speed is under-estimated. Overall, the wind speed is well reproduced at both scales, with $R^2 = 0.83$ globally and $R^2 = 0.89$ in the ROIs, and the point density well concentrated around the 1 : 1 line. The gauge corrected pressure is also well predicted at the global scale with $R^2 = 0.85$, but the ROI pressure shows the weakest fit with $R^2 = 0.45$. The sharp stagnation point and wake gradients around the structure are only partly recovered. The U-Net results (figure B1 in Appendix B) are comparable.

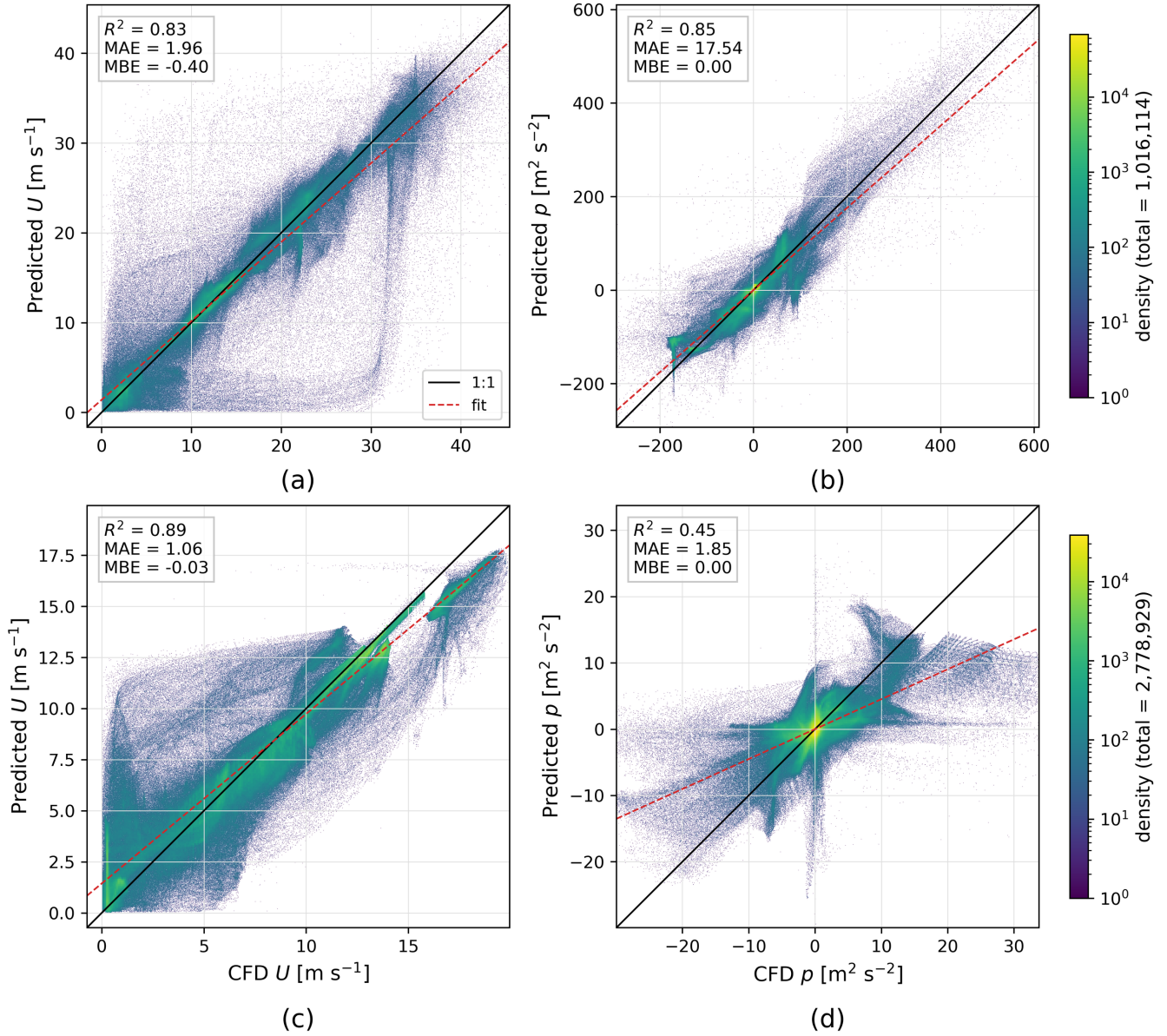


Figure 2. Predicted versus CFD density for the hybrid model over $> 1 \times 10^6$ points of the 18 test domains, for (a) global wind speed U ; (b) global gauge-corrected pressure p ; (c) ROI wind speed U ; (d) ROI gauge-corrected pressure p . The black line shows the 1:1 relation while the red line shows the linear fit. R^2 , MAE and MBE are given in each panel.

Finally, the surrogate inference is tested for speed. The hybrid model infers a complete terrain-scale field over a $3 \times 3 \text{ km}^2$ domain on the 30 m global grid in about 0.2 s and refines each structure ROI in about 1 s, so a full domain with a few structures is returned in a few seconds. The U-Net is heavier because of its tiled three-dimensional inference and it produces the terrain-scale field in 0.5 s and each ROI in about 10 s. These timings are measured on a mid-range laptop GPU (NVIDIA RTX 2000



395 Ada Generation). For comparison, a single OpenFOAM simulation of section 2.1 requires 2.5 to 4 hours on 36 CPU cores of
 an Intel Ice Lake cluster node.

3.2 Case-level analysis

The aggregate metrics reported above average over the 18 heterogeneous test domains. Resolving them per case identifies where
 the remaining error lives and makes the architecture trade-off concrete. We illustrate the behaviour with one representative case
 400 per domain family.

For complex terrain, we present the most difficult domain of the test set, a $3 \times 3 \text{ km}^2$ domain over a mountain pass, with
 global metrics above the average of the test set (the normalised RMSE for wind speed is 0.27, well above the 0.16 average
 for the hybrid model). This shows where the surrogate struggles most. Figure 3 shows the domain topography, the wind speed
 20 m above ground as simulated by OpenFOAM and as predicted by the final hybrid model, as well as a relative error map. The
 405 model reproduces the overall flow quite well. The acceleration of the wind over the crest on the right hand side, followed by a
 recirculation zone with low wind speed are well captured, while the left hand side also follows the general flow. The relative
 error map in figure 3 (d) localises where the errors are. The dominant patterns appear at the top-right crest and around it. Over
 the windward slope, the model over-estimates the wind speed, consistent with the systematic near-surface over-speed identified
 in section 3.1. The model over-predicts most strongly across the crest-to-wake transition, where the predicted transition from
 410 speed-up to separated flow is more gradual than in the reference. Within the recirculation zone, the wind speed is instead
 under-predicted and the predicted interior is more uniform than the partially recovering reference flow. Such paired errors of
 both signs show that the model reproduces the large-scale flow pattern but smooths the sharpest gradients, while the bulk of
 the domain remains largely unbiased. The difficulty of this domain therefore comes from the sharpness and the amplitude of
 its terrain-induced flow features rather than from a uniform model offset. We note that the U-Net prediction for this domain is
 415 comparable and thus not shown.

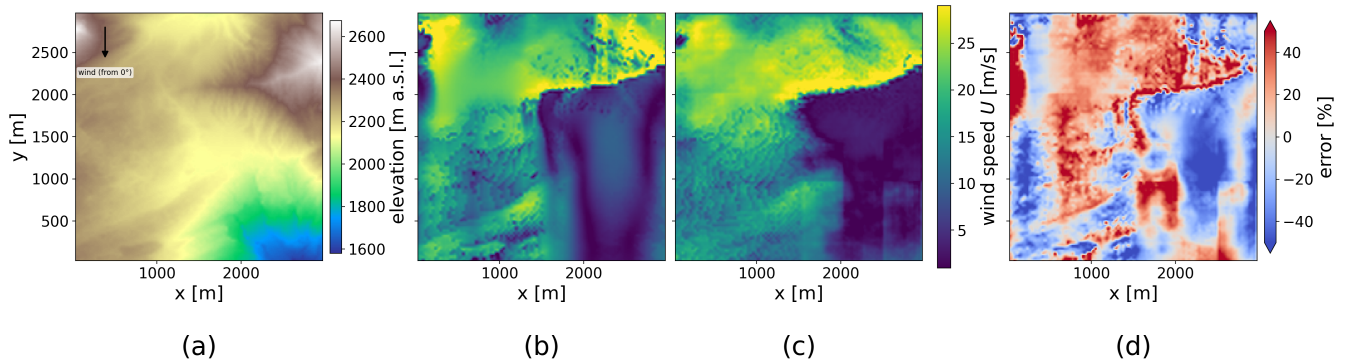


Figure 3. Complex terrain case. (a) terrain elevation with the inflow direction; (b) CFD reference wind speed at 20 m above ground; (c) Hybrid model prediction of wind speed at 20 m above ground; (d) relative error $(U_{\text{pred}} - U_{\text{CFD}})/U_{\text{ref}}$.



Figure 4 shows the wind speed 2.5 m above ground for the ROI of a test case with a single-structure (upper row, in this case a Helioplant®) and the ROI of a test case with a grid of structures (lower row, in this case 3×3 inclined panel grid). The fields obtained with the OpenFOAM simulations are presented on the left, while the predictions of the hybrid model are shown in the middle and the predictions of the U-Net are shown on the right.

420 For single-structures, shown in figure 4 (a-c), both models manage to capture isolated structures. The Stage 2 refiner adds the structure wake, represented by a low wind speed deficit downstream of the structure, on top of the background terrain flow. Both models reproduce this wake, but the hybrid model does so much more accurately, which is also seen in the metrics, with a normalised RMSE of 0.08 versus 0.16. Over all single-structures, this result is consistent, with an average normalised RMSE of 0.11 for the hybrid model and of 0.15 for the U-Net.

425 For the multi-structure array, shown in figure 4 (d-f), the result reveals an architecture trade-off. Indeed, the CFD data shows high wind speeds in the streamwise corridors between the structure columns, likely created by a channelling or Venturi effect. The U-Net model reproduces these inter-column channels rather well, whereas the hybrid model only produces a smoothed variant of the pattern. However, both models have an equivalent wind speed ROI error at 0.155. This is explained by the fact that the amount of cells affected by this high speed channelling is too small to strongly affect the ROI volume averaged

430 metric. Over all multi-structure cases, the U-Net performs slightly better, likely as a result of the architecture of the models. The U-Net applies three dimensional convolutions with receptive fields spanning the gap between neighbouring structures. The acceleration between structures thus naturally emerges from the convolutions. On the other hand, the hybrid model carries spatial context only in its two-dimensional terrain and structure encoders, which are sampled at each query point and passed to a point-wise head. Neighbouring output points are therefore not directly coupled, and the head can reproduce a gap jet only

435 if this sampled context already encodes the interaction between the surrounding structures. MLPs tend to favour smooth, low frequency functions. While Fourier features attempt to avoid this effect, the high speed jets are the most easily averaged away regions, especially if their smoothing does not strongly affect the loss function.

These results show that the best model is case-dependent. The hybrid model offers the best aggregate accuracy and flexible point-wise inference, and is also faster. The U-Net resolves more faithfully inter-structure interactions that can matter for

440 engineering applications such as structure loadings.

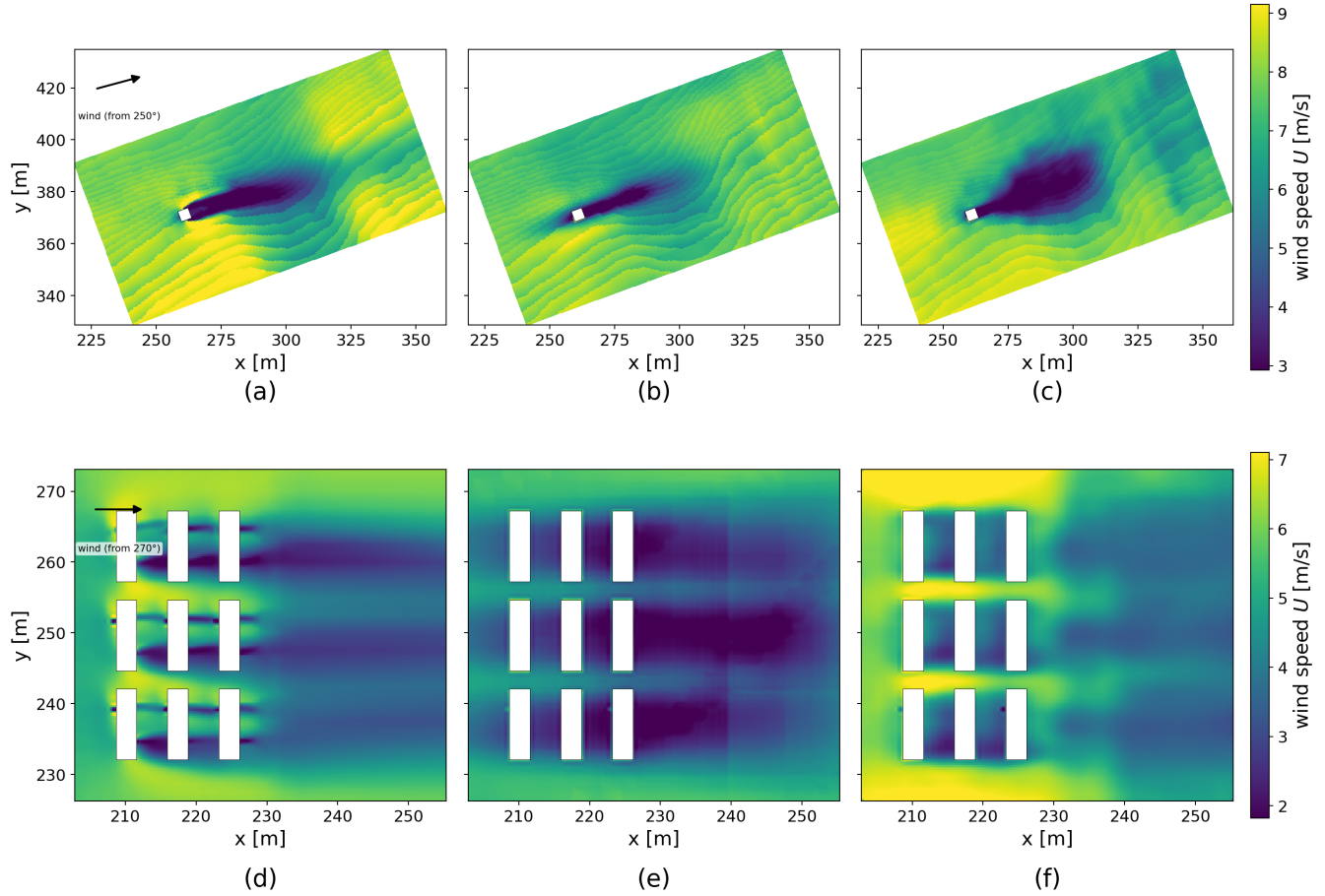


Figure 4. Structure cases with structure footprints in white. (a) CFD reference wind speed at 2.5 m above ground for a single-structure case; (b) Hybrid model prediction of wind speed at 2.5 m above ground for a single-structure case; (c) U-Net model prediction of wind speed at 2.5 m above ground for a single-structure case; (d) CFD reference wind speed at 2.5 m above ground for a multi-structure case; (e) Hybrid model prediction of wind speed at 2.5 m above ground for a multi-structure case; (f) U-Net model prediction of wind speed at 2.5 m above ground for a multi-structure case

3.3 Contribution of the physics losses

While supervised metrics have been reported above, additional, unsupervised values can also be assessed. We thus introduce the continuity and the momentum root mean square (RMS) residuals, given by:

$$\text{RMS}_{\text{div}} = \sqrt{\frac{1}{M} \sum_{j=1}^M \left(\frac{\nabla \cdot \mathbf{u}_j}{s_{\text{div}}} \right)^2}, \quad \text{RMS}_{\text{mom}} = \sqrt{\frac{1}{M} \sum_{j=1}^M \frac{\|\mathbf{r}_j\|^2}{s_{\text{mom}}^2}} \quad (10)$$

that is, the RMS of the same normalised residual operators as in eqs. (2)-(3), which measure how well a prediction satisfies the governing equations independently of the reference data. RMS_{mom} uses the effective viscosity of the evaluated domain, with



ν_t taken from its reference simulation, as it is not yet a direct output of the model. It measures the consistency of the predicted velocity and pressure with the momentum balance of the turbulence closure that generated the reference. On a domain without CFD wind field, only a constant-viscosity ($\nu = 1.5 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$) momentum residual can be formed. The conclusions drawn below do not depend on the assumed ν_t , as evaluating the same predicted field with constant viscosity only, using no reference information beyond the model prediction, gives the same contrast between the physics-regularised and the data-only models.

As mentioned in section 2, isolating the contribution of the physics losses is done by training each stage as a matched pair that differs only in the differential-equation weights. This comparison is run on the final 292 training domains without streamwise-mirror augmentation, for both model families and both stages. Since both members of each matched pair share an identical training setup, the augmentation setting should not strongly affect the isolation of the physics terms. The effect of the augmentation itself was quantified separately in section 3.1. Table 4 first compares the models evaluated on Stage 1, so on the global 30 m resolution domains. The two model families reach similar global wind speed and pressure accuracy, with normalised RMSE values between 0.17 and 0.19 for U and between 0.26 and 0.33 for p . Adding physics losses to the training leaves the wind speed error within the seed range, while the global pressure error decreases by 10% (hybrid) and 24% (U-Net). The continuity and momentum residuals are systematically lower with physics included in training, but only modestly. This suggests that, at this scale, the terrain flow is generally smooth and well supervised, resulting in a reasonably good physics conservation. It can however be noted that, for the U-Net model, the momentum residual is tripled without physics losses in training.

To separate the contribution of CFD error and grid resampling from that of the surrogate, the same residual operators were also applied to the reference CFD fields themselves, on the same grids over the 18 test domains. Table 4 reports these values. The CFD solution returns non-zero yet small residuals of 0.004 for both continuity and momentum. The model remains well above these values, by a factor of 3 to 17, showing that the reported residuals are dominated by the surrogate error and that comparisons between physics-regularised and data-only models are meaningful.

Table 4. Test normalised RMSE for the global wind speed and pressure, normalised continuity residual RMS and momentum residual RMS, evaluated on the 30 m resolution global domain for physics-regularised versus data-only models. The last row gives the residual obtained by applying the same operator to the reference CFD fields resampled on the same grids.

Model	training	global U	global p	div res.	mom res.
Hybrid	with physics	0.172	0.264	0.016	0.029
Hybrid	data-only	0.167	0.294	0.019	0.038
U-Net	with physics	0.184	0.249	0.011	0.023
U-Net	data-only	0.192	0.328	0.017	0.069
CFD ref.	-	-	-	0.004	0.004

Table 5 shows the same metrics for each model evaluated on Stage 2, i.e. the 0.5 m ROIs. Again, both families reach close normalised RMSE values for wind speed and pressure. The hybrid model attains lower ROI pressure errors, while the U-Net



reaches lower ROI wind speed errors. The physics losses have a clear effect on the residuals: for both model families, the residuals are almost twice as high with the data-only training loss. Table 5 also shows the CFD fields residuals on the ROI grids. Again, the reference residuals are small non-zero values of 0.018 and 0.028 for continuity and momentum respectively, smaller than the surrogates' residuals by a factor of 9 to 26.

Table 5. Test normalised RMSE for the ROI wind speed and pressure, normalised continuity residual RMS and momentum residual RMS, evaluated on the 0.5 m resolution ROIs for physics-regularised versus data-only models. The last row gives the residual obtained by applying the same operator to the reference CFD fields resampled on the same grids.

Model	training	ROI U	ROI p	div res.	mom res.
Hybrid	with physics	0.203	0.107	0.166	0.277
Hybrid	data-only	0.213	0.121	0.322	0.459
U-Net	with physics	0.186	0.238	0.238	0.428
U-Net	data-only	0.209	0.221	0.466	0.730
CFD ref.	-	-	-	0.018	0.028

Overall, the principal contribution of the physics losses is not a large reduction of the aggregate wind speed error, which is already strongly constrained by the supervised data. Their main effect is to improve the consistency of the predicted fields with mass and momentum conservation. This behaviour is consistent with reported difficulties of jointly optimising data and physics objectives, where a Pareto front between fitting the reference data and satisfying the governing equations can emerge (Rohrhofer et al., 2023). Treating the residuals as physically motivated regularisers, evaluated separately from the supervised error, is coherent with this observation. We note that in the development experiments at the initial 120-domain budget (Appendix A), adding the two physics residuals to a purely data-driven MSE baseline reduced the ROI pressure error from 0.434 to 0.164. Although these development runs were not matched pairs, this suggests that the physics losses improve the supervised accuracy most when training data is scarce, and that their contribution shifts to physical consistency as the data budget grows.

3.4 Assessment of the two-stage cascade

Both model families share the two-stage cascade introduced in section 2.2, in which a terrain-scale background prediction is refined locally around the structures. Table 6 quantifies the benefit of this design by comparing both final families after Stage 1 alone and after adding the Stage 2 refiner, on the final 292-domain dataset. The Stage 1 rows evaluate the background, global predicted fields inside the ROIs, without any refiner. Both backgrounds reach comparable global accuracy, but neither resolves the structure-scale flow on its own. This is further illustrated in figure 5 for a multi-structure case. The Stage 1 background (b) indeed produces a smooth flow over the entire ROI, whereas the Stage 2 refiner (c) adds the metre-scale wake behind the structures on top of that background, more accurately approaching the reference CFD data (a). This is most effective for the U-Net, whose background reaches high ROI errors. Adding the Stage 2 refiner has a large impact on the U-Net, reducing these large errors from 0.263 to 0.176 for wind speed and from 0.550 to 0.130 for pressure. The hybrid background is already much



495 better in the ROIs, particularly for the pressure, but the refiner still lowers its ROI wind speed error from 0.24 to 0.17. The two-stage cascade is therefore beneficial for both families, and essential for the U-Net whose global background alone cannot resolve the structure-induced variations.

Table 6. Test normalised RMSE of wind speed and pressure on the Stage 1 global 30 m grid and the Stage 2 0.5 m ROIs, with the validation score (eq. (8)) for cascade architecture comparison on the 292-domain training set.

Model	selector	global U	global p	ROI U	ROI p
Hybrid, background (Stage 1)	0.357	0.157	0.247	0.236	0.132
Hybrid + refiner (Stage 2)	0.338	0.157	0.247	0.169	0.122
U-Net, background (Stage 1)	0.426	0.177	0.233	0.263	0.550
U-Net + refiner (Stage 2)	0.365	0.177	0.233	0.176	0.130

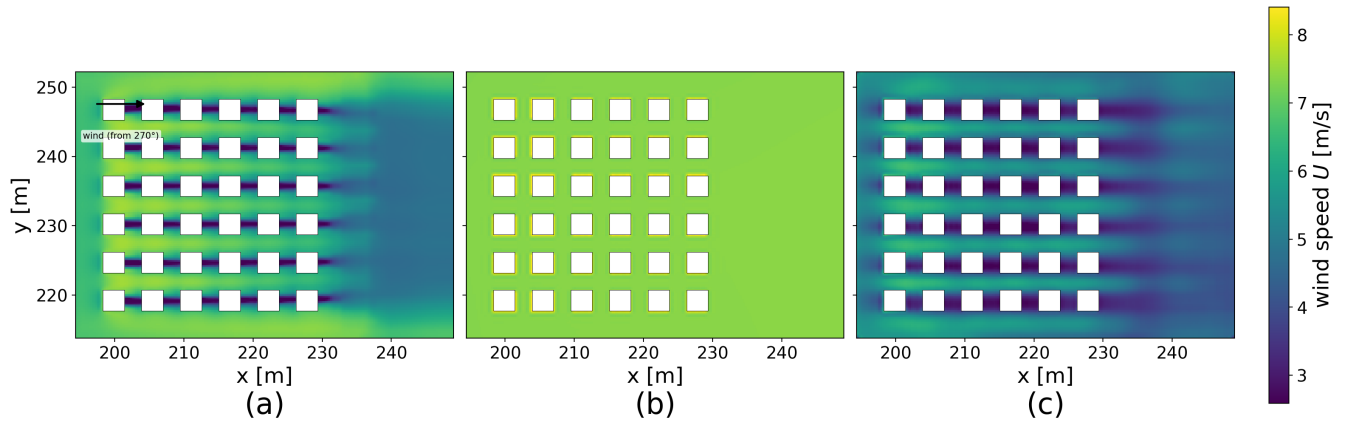


Figure 5. Benefit of the cascade on a multi-structure test domain (a) CFD reference wind speed at 2.5 m above ground; (b) U-Net Stage 1 prediction at 2.5 m above ground; (c) U-Net Stage 2 prediction at 2.5 m above ground.

3.5 Active learning

Active learning is used to expand the training set in rank order of an acquisition score combining continuity and momentum residuals, boundary-condition error and feature-space diversity, evaluated on large pools of unsimulated candidate domains. The methodology follows Rygiel et al. (2025). A typical iteration scores a pool of 500 randomly generated candidate domains and retains the 20 highest. Low-scoring candidates are consistently the "easy" ones, such as flat terrain sampled over a lake, whereas high-scoring candidates are the steep, separation-prone topographies. The two model scales (global and ROIs) are targeted in active learning in two successive phases. Each phase was run with the model family best suited to it, the terrain-scale background with the hybrid, the structure-scale refiner with the U-Net.

505 The first phase grew the training dataset from 120 to 194 domains with complex-terrain cases, developing the hybrid terrain-scale background, through different iterations of active learning. Table 7 (upper block) reports the wind speed and pressure



normalised RMSE on the test set as well as the validation selector for some iterations. The global errors are found to strongly decrease from 0.26 to 0.18 for the wind speed and from 0.49 to 0.28 for pressure. The bias is found to improve alongside the RMSE through these iterations. At 120 domains the near-ground signed wind speed bias reached +0.21 and the global pressure absolute bias 0.29. Both fall by roughly a factor of two to three by the 194-domain iteration. As none of these additional training domains contained structures (thus ROIs), the ROI errors do not show a consistent trend with additional data. In this phase, score-based selection is compared to random selection at a matched budget of 130 domains (so adding 10 domains). Score-ranked selection reaches a lower validation score (0.491 versus 0.520) and a lower test global wind speed error (0.228 versus 0.253) than random selection. While individual differences may be close to the seed range, they are generally outside of it, and the effect is consistent as the selected set is better on five out of five reported metrics, in both validation and test splits. This supports a consistent, moderate benefit of score-based selection at the terrain scale.

In the second phase, the training dataset is augmented from 194 to 292 domains with cases including structures. To isolate the data efficiency of the Stage 2 refiner, its global 30 m scale background is frozen to the 194-domain U-Net and only the ROI refiner is retrained on increasing amounts of selected structure data. Table 7 lower block shows the errors associated with different iterations of this second phase of active learning. The ROI-scale errors decrease with additional structure domains, but the effect is small. The wind speed error stays around 0.21, while the pressure error goes from 0.19 to 0.17. Separating by structure families reveals that the single-structure wind speed error is flat at around 0.16 and the multi-structure error improves slightly from 0.28 to 0.26. Furthermore, the validation score is itself non-monotonic. This second phase active learning targeted at structures is thus less efficient than the terrain phase, as most of the structure-scale accuracy is already reached with the original structure set, and the remaining multi-structure error appears limited by the frozen 194-domain background rather than by the amount of structure data. Consistently, the fully retrained two-stage models, in which the background is retrained jointly with the refiner on all 292 domains, as for the final models, recover most of the multi-structure gap, lowering the U-Net multi-structure ROI wind speed from 0.26 on the frozen-background curve to 0.21.



Table 7. Test normalised RMSE of wind speed and pressure on the Stage 1 global 30 m grid and the Stage 2 0.5 m ROIs, with the validation score (eq. (8)) for Active Learning iterations. Upper block: terrain-scale phase using the hybrid model, growing the set from 120 to 194 complex-terrain domains. Lower block: structure-scale phase using the U-Net ROI refiner on a frozen 194-domain background, growing the structure set to 292 domains.

Training set	selector	global U	global p	ROI U	ROI p
<i>Phase 1 - terrain-scale (hybrid background)</i>					
120 domains (initial)	0.522	0.259	0.491	0.267	0.148
130 domains (AL)	0.491	0.228	0.392	0.258	0.147
158 domains (AL)	0.456	0.195	0.360	0.268	0.125
194 domains (AL)	0.449	0.177	0.275	0.268	0.132
<i>Phase 2 - structure-scale (U-Net refiner, frozen 194 background)</i>					
194 (baseline)	0.411	0.188	0.267	0.215	0.192
+40 structures (234)	0.436	0.188	0.267	0.211	0.193
+98 structures (292)	0.398	0.188	0.267	0.207	0.178

3.6 Limitations

Several limitations need to be considered when assessing the validity of the surrogate. First, the reference solutions are steady-state RANS fields with the standard $k-\varepsilon$ closure, which is known to under-predict separation and recirculation, and the steady formulation represents intrinsically unsteady wakes by their stationary mean. The surrogate inherits these modelling choices by construction, and its errors must be interpreted relative to this RANS reference rather than to the true atmospheric flow. Furthermore, all results are obtained on domains drawn from the same generator and evaluated against the same $k-\varepsilon$ RANS setup. Comparison against independent references such as a large eddy simulation, a wind tunnel configuration or a field campaign, is required before the surrogate is fully validated.

Second, the dataset spans a deliberately restricted set of atmospheric conditions. The inflow is an idealised, neutral logarithmic profile with a fixed aerodynamic roughness length z_0 , uniform over the domain. Real inflows deviate from this idealised shape, with wind speed and direction varying with height under the influence of upstream terrain, and the surface roughness varies with land cover by orders of magnitude, affecting the near-ground wind field. These simplifications isolate the terrain- and structure-induced responses, which is the central object of this study. The fixed neutral stratification is a further restriction. Extending the model towards more realistic input profiles and spatially varying roughness will require additional CFD cases covering such cases.

Third, the wind turbines are represented as static, parked geometries. These still produce a wake through their drag, but this wake differs from that of an operating rotor, whose velocity deficit is governed by the thrust and energy extraction at the operating point and which additionally contains rotation-induced swirl. The turbine cases therefore validate the surrogate's ability to handle large slender structures, not operational wake modelling. Representing operating rotors, for example with



OpenFOAM actuator disc models, should be considered in a future version. Moreover, the true capability of the surrogates to accurately generalise to different structures remains currently under-tested, with 10 different structures used in the training set and only 7 in the test set, all of which appear in training, so unseen geometries are not tested.

Fourth, the current surrogate predicts the mean velocity and kinematic pressure fields but does not predict turbulence quantities. Each reference CFD simulation already provides the turbulent kinetic energy (k), its dissipation rate (ε) and the turbulent kinematic viscosity (ν_t). These fields are currently used only indirectly, through ν_t in the momentum residual. Because ν_t is taken from the reference simulation, it is available for the simulated training, validation and test domains but not for other unsimulated domains, so the physical consistency of a prediction on a new domain can only be assessed through the continuity residual and a constant-viscosity momentum residual. Extending the network outputs to include (k, ε, ν_t) or related turbulence quantities is therefore a direct next step and would allow the surrogate to provide information relevant to turbulence intensity, fatigue loading and wake recovery, as well as a self-contained momentum residual on new domains. Such an extension will require suitable output scaling, additional supervised and physical losses, and dedicated evaluation metrics, but does not require regeneration of the existing CFD database.

Fifth, the ROI pressure is the least accurate output ($R^2 = 0.45$, see figure 2), and the sharp stagnation and wake gradients around structures, which are particularly relevant for loading features, are only partially captured. Structural load estimates derived from the predicted pressure should therefore be used as preliminary information between candidate layouts rather than substitute for load calculations. Published wind-tunnel measurements of wind loading on ground-mounted panels and arrays (Zhao et al., 2025; Abiola-Ogedengbe et al., 2015; Stathopoulos et al., 2014) provide reference pressure and force coefficients against which such load proxies can be evaluated.

Finally, neither architecture fully resolves all relevant flow features. The hybrid model gives the best aggregate accuracy and fastest, most flexible inference, but can smooth localised inter-structure accelerations, which may be important for snow accumulation, for example. The U-Net better captures some spatial interactions between neighbouring structures, but is heavier and less accurate in several aggregate metrics. Further development should therefore target both improved spatial coupling in the hybrid model and better accuracy and efficiency in the U-Net. Future development will address these limitations.

4 Conclusions

This work presented a two-stage cascade CFD surrogate for wind and pressure prediction over complex terrain and structures. The model connects two spatial regimes within one framework. In Stage 1, a global background field over domains of up to $3 \times 3 \text{ km}^2$ with a resolution of 30 m and in Stage 2 a structure refiner over regions of interest with a resolution of 0.5 m. Two architecture families were developed and compared, with a flexible point-wise hybrid MLP and a grid-based U-Net.

On the frozen test set, the final models predict the global scale and the ROI scale flow with comparable accuracy. The hybrid MLP reaches a global wind speed normalised RMSE of 0.157 and an ROI wind speed normalised RMSE of 0.169, while the U-Net reaches just below 0.18 for both. Both models have a noticeably higher pressure normalised RMSE with 0.247 for the hybrid MLP and 0.233 for the U-Net, brought down to 0.202 and 0.185 after gauge correction. Both models are found to



over-predict wind speed in the near-ground region. The direction error is small in the bulk of the domains and a little higher near-ground. Pooled over all test points, both surrogates explain most of the variance ($R^2 = 0.83 - 0.89$) for the wind speed at both scales and for the pressure at terrain-scale, while the structure-scale pressure is the hardest field to reproduce. Inference requires seconds per domain on a mid-range laptop GPU instead of hours for a new CFD simulation. The hybrid MLP provides the fastest and most flexible predictions, while the U-Net provides stronger spatial coupling and better resolves the accelerated jets between closely spaced structures, which may be important for certain applications.

The results address all questions posed in the introduction. First, a single surrogate can predict both kilometre-scale and metre-scale structure-induced with normalised wind speed errors of 0.16 to 0.18 with respect to the reference RANS, at low computational cost. The remaining errors are concentrated in the most difficult regions and cases. This concerns the near-ground layer, where the models tend to over-predict the wind speed, and the sharpest flow gradients, such as crest speed-ups and transitions into recirculation zones in steep, separation-prone domains. At the structure scale, it particularly concerns inter-structure accelerations of dense arrays. Second, the physics losses contribute mainly as a physical regulariser rather than as a direct driver of lower supervised error. The physics-regularised models indeed do not show reduced wind speed RMSEs compared with matched data-only models and only a moderate reduction of the pressure error, but they produce fields that satisfy mass and momentum conservation more closely. This is attributed to the fact that we have sufficient CFD data for training, while the development experiments suggest that the physics losses help more when less training data is available, although this was not tested as matched pairs. Third, the results show that the two-stage cascade is an important model design. A global, 30 m resolution field alone is not sufficient to resolve local structure-scale flow. The reduction in ROI error with a Stage 2 refiner is particularly large with the U-Net, but no model is able to recover effects such as structure wakes without the refiner. The two architecture families represent a trade-off. The hybrid gives the best aggregate accuracy, fastest inference and arbitrary coordinate evaluation, while the U-Net resolves some inter-structure accelerations that tend to be smoothed by the hybrid. As of now, the best choice is therefore case- and application-dependent. Both families are still kept for future developments. Fourth, the expansion of the training set was most effective during the complex terrain phase, and score-based selection outperformed random selection at a matched budget. Selected terrain cases strongly improved the global wind speed and pressure errors over successive iterations. The structure-scale Stage 2 seems to be less data-efficient as a large part of the accuracy is already reached with the original dataset. In future developments, active learning will continue to be used.

Several limitations still remain. The current work covers a restricted physical type of flow with neutral stratification and wind turbines are represented as stationary geometries. Future datasets could address these by using for example buoyant-BoussinesqSimpleFoam instead of simpleFoam, which handles temperature-driven buoyancy, as well as actuator disks for wind turbines. A direct extension of the model is to predict turbulence quantities in addition to wind speed and pressure. Each CFD simulation already provides the turbulent kinetic energy, dissipation rate and turbulent kinematic viscosity. These fields are currently unused, except for the latter which is only used for momentum residual calculation. Adding them as outputs would provide information on turbulence intensity or other important engineering features without requiring additional data. While the different developments, mainly the weighted Charbonnier loss formulation with physics residuals, the two-stage cascade, the active learning expansion and the streamwise-mirror augmentation, have already brought the error down from the



initial data-only MSE hybrid at 120 domains to the final models, from 0.34 to 0.16 for the global wind speed and from 0.43 to 0.12 for the ROI pressure, further work should target the over-estimation of the near-ground wind speed, representation of inter-structure acceleration in the hybrid model and single-structure sharpness for the U-Net model. Validation of the surrogate against field measurements in complex terrain, for example using recent LiDAR campaigns should also be considered, and a more in depth assessment of the capability of the model to generalise flow around structures should be conducted.

Overall, the results demonstrate that a physics-regularised, actively expanded CFD surrogate can connect terrain-scale atmospheric flow and metre-scale structure effects in a single fast prediction framework, enabling fast preliminary siting, layout comparison and structure load screening.

Code and data availability. The source code of the surrogate pipeline and the trained model weights are publicly available at <https://github.com/jimmygasser/pinnfluid>. An interactive application built on the final models, allowing terrain selection, structure placement and prediction over Swiss domains, is publicly accessible at <https://www.pinnfluid.ch>. The CFD training dataset (292 simulated domains, resampled fields) is archived on EnviDat under <https://doi.org/10.16904/envidat.810> (Gasser and Lehning, 2026).

Appendix A: Model development experiments

This section details the development experiments summarised in section 2.4. All experiments reported here were performed with the hybrid model family on the initial 120-domain training set, without the streamwise-mirror augmentation, and guided by the validation score of eq. (8). The validation and test sets are the frozen 18-domain splits used throughout the paper. These development runs were performed with the development configuration and data version. Their absolute values therefore differ from those of the models retrained with the final configuration in Table 7, and only the relative changes between the configurations of this appendix should be read.

First, the effects of the supervised and the physics losses were analysed. In contrast to section 3.3, which switches the physics terms of the final loss formulation on and off in matched pairs, the experiments here trace how the loss formulation itself was constructed. Starting from a purely data-driven MSE baseline, the continuity residual, the momentum residual, the Charbonnier loss and the spatial weighting were added incrementally. Table A1 shows the test metrics and the validation score for the different loss combinations at fixed model parameters, isolating the contribution of the loss formulation. It can be seen that adding continuity to a purely data driven MSE model leads to strong error reduction on the Stage 2 ROI domains, while not hurting Stage 1 global metrics. Adding momentum reinforces this effect and even decreases global Stage 1 metrics. With both physics residuals included, the pressure error in ROIs is for example reduced from 0.434 to 0.164. As shown in section 3.3, on the larger 292-domain training dataset, the effect of the physics losses on the supervised accuracy of the model is more limited. This suggests that physics losses are most useful for improving accuracy when the training budget is small, while at larger data budgets their main contribution becomes the improved physical consistency of the predicted fields. Replacing the MSE supervised loss by a Charbonnier loss mostly improved Stage 2 metrics. Overall, a Charbonnier formulation with spatial



weighting produces the best validation score and the best metrics on all but one score. Table A1 also indicates that the global wind field is strongly constrained by the supervised data, whereas the physics residuals and robust supervised loss contribute most to the pressure prediction and to the structure-scale regions.

Table A1. Test normalised RMSE of wind speed and pressure on the Stage 1 global 30 m grid and the Stage 2 0.5 m ROIs, with the validation score (eq. (8)) for different loss/physics combinations at fixed data (120 training domains) and architecture (hybrid).

Loss configuration	selector	global U	global p	ROI U	ROI p
Data-only (MSE)	0.636	0.340	0.675	0.315	0.434
+ continuity residual	0.606	0.341	0.676	0.303	0.211
+ momentum residual	0.600	0.336	0.655	0.346	0.164
Charbonnier + physics	0.581	0.335	0.664	0.263	0.156
Weighted Charbonnier + physics	0.545	0.333	0.633	0.257	0.130



650 The main architectural and training hyperparameters were then screened one factor at a time around a baseline configuration. Figure A1 summarises the complete screening of about 40 configurations, grouped into (a) architecture variations, (b) training-data sampling variations and (c) supervised-loss and physics variations. The results show that the prediction is generally weakly sensitive to architectural hyperparameter changes, as varying the neural network width, depth, the two-dimensional encoder width, the number of Fourier features or their bandwidth changes the normalised RMSE by a few percent only (figure A1 (a)).

655 This is particularly true for the Stage 1 global field, while some stronger sensitivities are found on ROIs. For the data sampling (b), increasing the number of patches does not systematically improve performance and can degrade the ROI velocity error, whereas reweighting the supervision towards the regions of interest and the near-structure wakes lowers the ROI pressure error. The loss variations (c) are consistent with Table A1, with the weighted Charbonnier formulations reaching the best selector scores and the physics-free variant the worst ROI pressure. The purpose of this screening was to select a configuration for

660 subsequent development and not to rank variants, as multiple error differences are comparable to the seed range. The main outcome is that the final accuracy is not strongly sensitive to most of these parameters, which motivated the choice to retain an architecture close to the baseline and to concentrate improvements on the loss formulation and the training data.

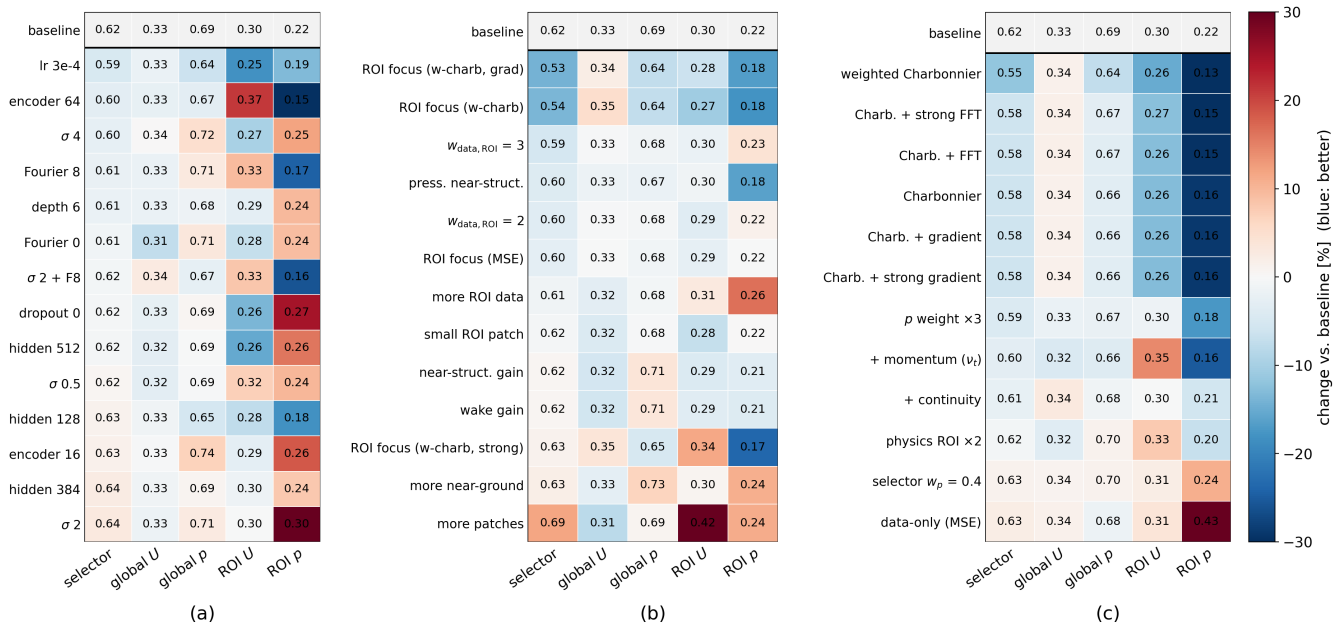


Figure A1. Development screening at 120 training domains. Each cell reports the test normalised RMSE and validation selector of one configuration; colour encodes the relative change with respect to the baseline configuration (blue: better). (a) one-factor-at-a-time variations of the architecture; (b) variations of the training-data sampling and (c) variations of the supervised loss and physics.



Appendix B: Supplementary figures

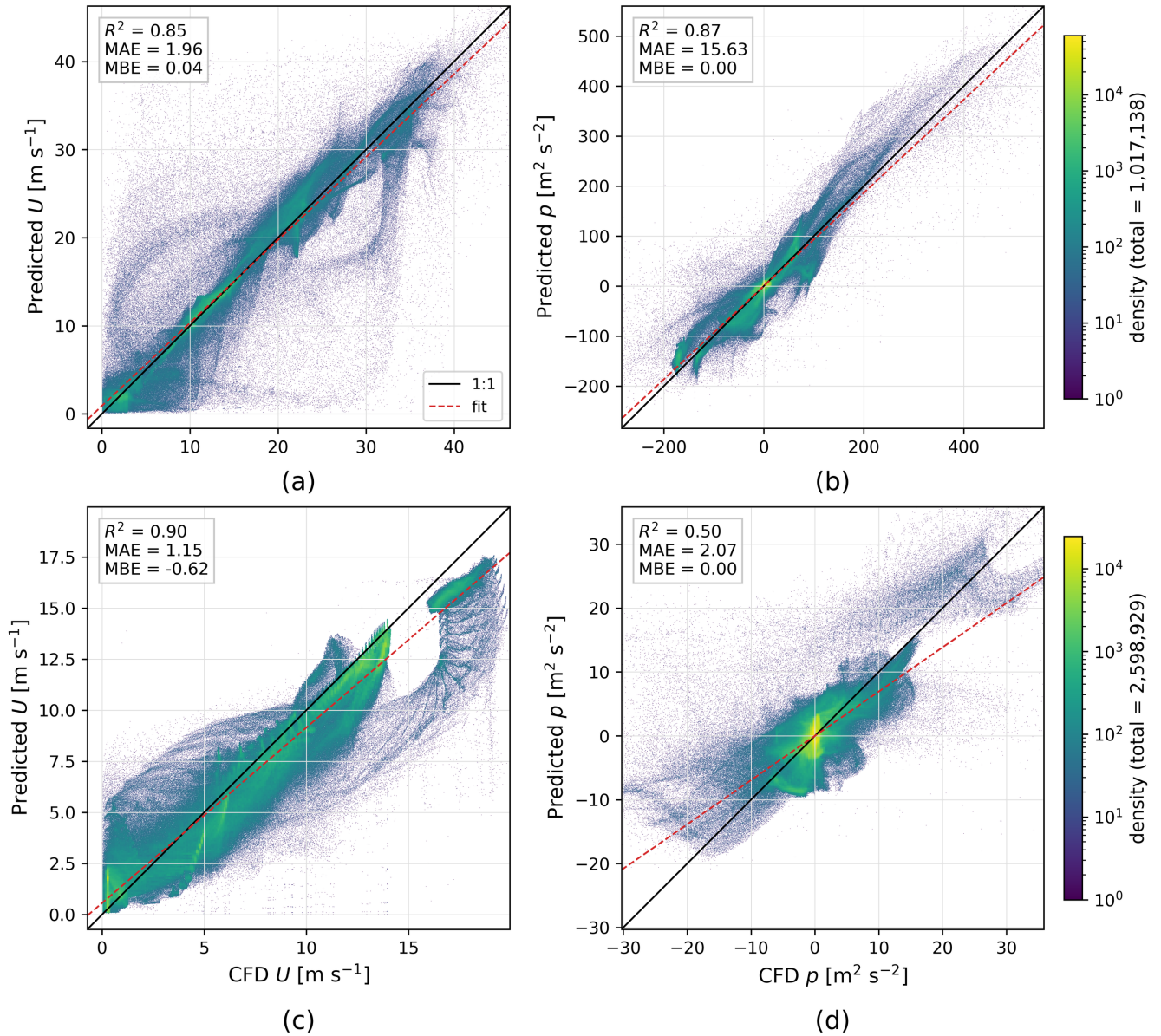


Figure B1. Predicted versus CFD density for the U-Net model over $> 1 \times 10^6$ points of the 18 test domains, for (a) global wind speed U ; (b) global gauge-corrected pressure p ; (c) ROI wind speed U ; (d) ROI gauge-corrected pressure p . The black line shows the 1:1 relation while the red line shows the linear fit. R^2 , MAE and MBE are given in each panel.

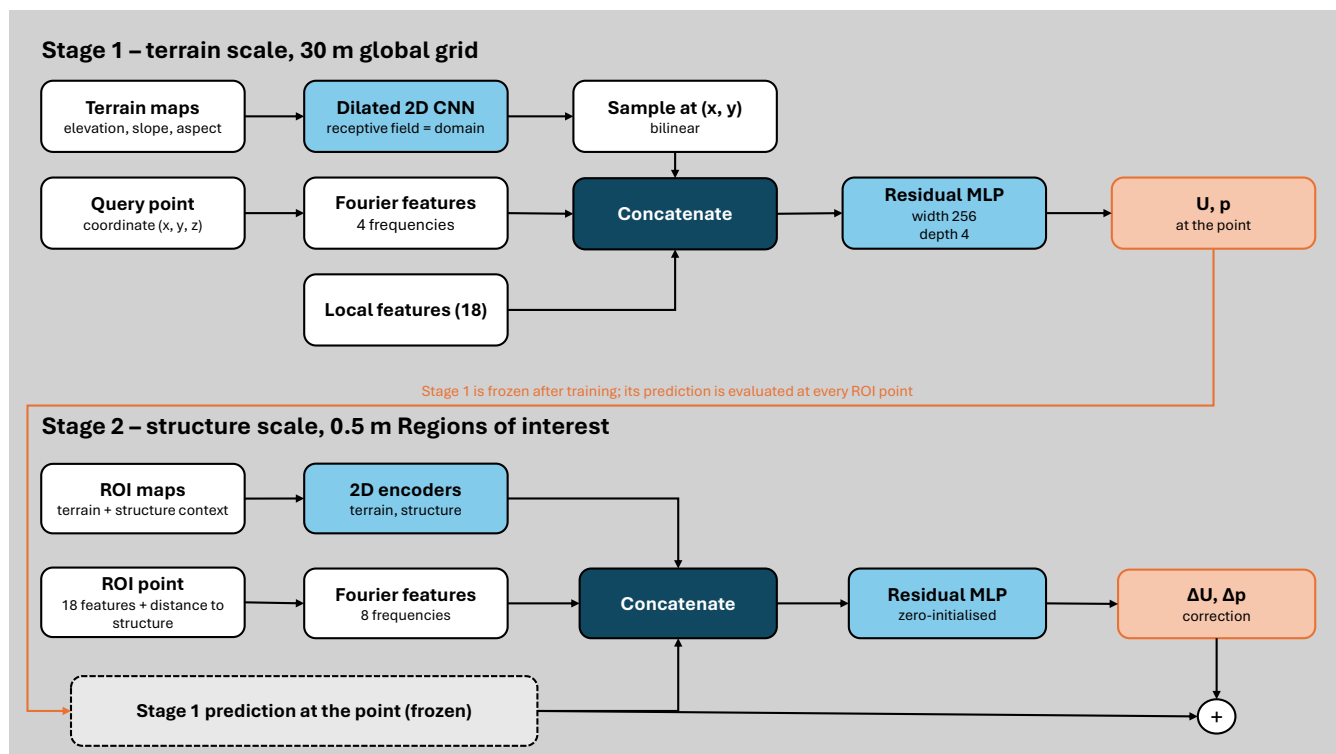


Figure B2. Architecture of the hybrid model family. Stage 1 (top) predicts the terrain-scale background from local features and the encoded terrain map; Stage 2 (bottom) predicts a residual correction on the 0.5 m ROIs from the ROI features, the structure context and the frozen Stage 1 prediction evaluated at every ROI point.

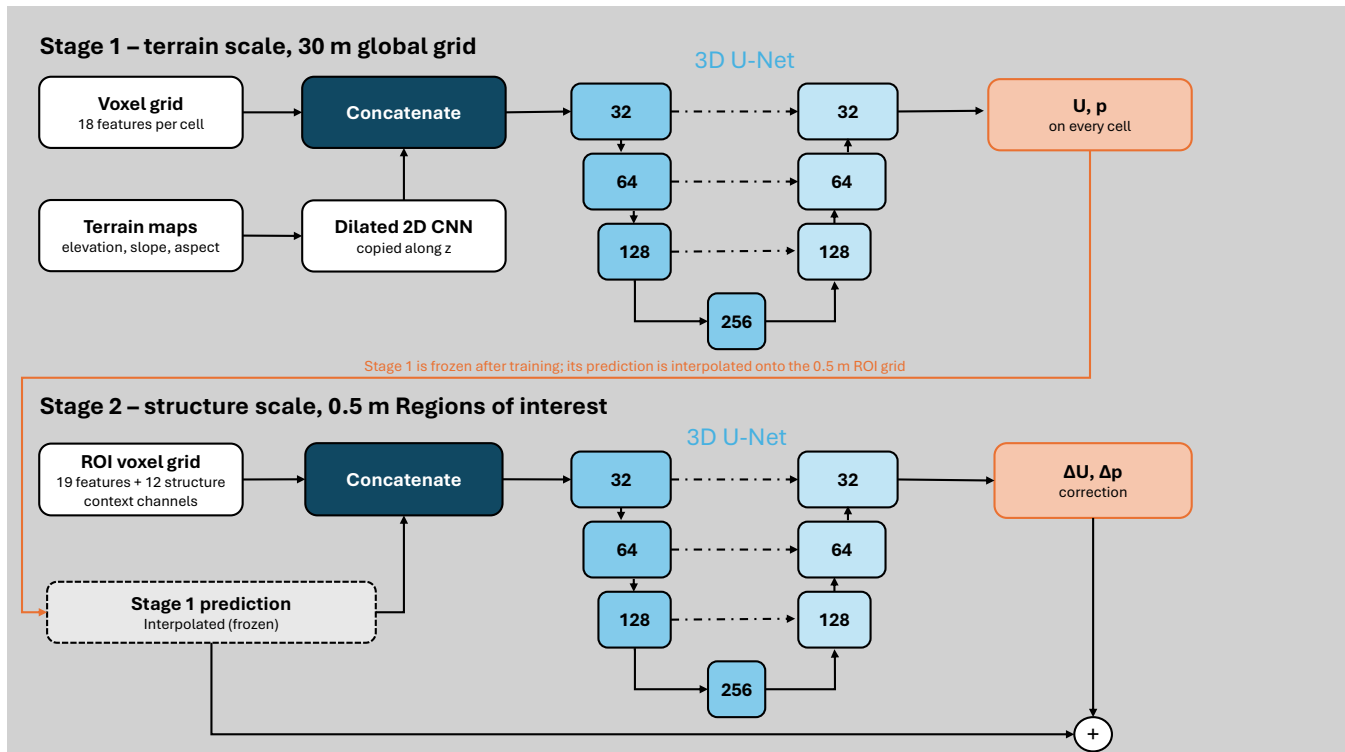


Figure B3. Architecture of the U-Net model family. Stage 1 (top) predicts the terrain-scale background on the 30 m voxel grid; Stage 2 (bottom) is a second 3D U-Net that predicts a residual correction on the 0.5 m ROI grid from the ROI features, the structure context and the frozen Stage 1 prediction interpolated onto the ROI grid.



Appendix C: Supplementary equations

665 The boundary-condition terms of eq. (4) are, explicitly,

$$\begin{aligned}\mathcal{L}_{\text{inlet}} &= \frac{1}{|\mathcal{B}_{\text{in}}|} \sum_{i \in \mathcal{B}_{\text{in}}} \frac{\|\mathbf{u}_i - \mathbf{u}_i^{\text{log}}\|^2}{U_{\text{ref}}^2}, & \mathcal{L}_{\text{outlet}} &= \frac{1}{|\mathcal{B}_{\text{out}}|} \sum_{i \in \mathcal{B}_{\text{out}}} \frac{(p_i - p_{\text{ref}})^2}{s_p^2}, \\ \mathcal{L}_{\text{side}} &= \frac{1}{|\mathcal{B}_{\text{side}}|} \sum_{i \in \mathcal{B}_{\text{side}}} \frac{(\mathbf{u}_i \cdot \mathbf{n}_i)^2}{U_{\text{ref}}^2}, & \mathcal{L}_{\text{top}} &= \frac{1}{|\mathcal{B}_{\text{top}}|} \sum_{i \in \mathcal{B}_{\text{top}}} \frac{(\mathbf{u}_i \cdot \mathbf{n}_i)^2}{U_{\text{ref}}^2},\end{aligned}$$

where $\mathbf{u}^{\text{log}}$ is the analytic logarithmic inflow profile, $p_{\text{ref}} = 0$ the outlet reference, $\mathbf{n}$ the outward boundary normal, s_p a kinematic-pressure scale, and $\mathcal{B}_{\text{in}}, \mathcal{B}_{\text{out}}, \mathcal{B}_{\text{side}}, \mathcal{B}_{\text{top}}$ the sets of sampled points on the corresponding boundary faces.

670 *Author contributions.* JG: conceptualisation, methodology, software, investigation, data curation, formal analysis, visualisation, writing (original draft). YF: conceptualisation, resources, writing (review and editing). ML: conceptualisation, supervision, funding acquisition, writing (review and editing).

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The CFD simulations were performed on the Jed cluster of the EPFL Scientific IT and Application Support (SCITAS),
 675 and the machine learning models were trained on the GPU cluster of the EPFL Research Computing Platform (RCP). The authors thank the Swiss Federal Office of Topography swisstopo for providing the swissALTI3D digital elevation model. Some of the authors are also part of the Speed2Zero and UrbanTwin consortia, which are projects financed by the ETH Board.

Financial support. This research has been supported by the Swiss Federal Office of Energy (SFOE) as part of the SWEET consortium EDGE. The authors bear sole responsibility for the conclusions and the results presented in this publication.



680 References

- Abiola-Ogedengbe, A., Hangan, H., and Siddiqui, K.: Experimental investigation of wind effects on a standalone photovoltaic (PV) module, *Renewable Energy*, 78, 657–665, <https://doi.org/10.1016/j.renene.2015.01.037>, 2015.
- Achermann, F., Stastny, T., Danciu, B., Kolobov, A., Chung, J. J., Siegwart, R., and Lawrance, N.: WindSeer: real-time volumetric wind prediction over complex terrain aboard a small uncrewed aerial vehicle, *Nature Communications*, 15, 3507, <https://doi.org/10.1038/s41467-024-47778-4>, 2024.
- 685 Aly, A. M. and Clarke, J.: Wind design of solar panels for resilient and green communities: CFD with machine learning, *Sustainable Cities and Society*, 94, 104 529, <https://doi.org/10.1016/j.scs.2023.104529>, 2023.
- Berg, J., Mann, J., Bechmann, A., Courtney, M., and Jørgensen, H.: The Bolund Experiment, Part I: Flow Over a Steep, Three-Dimensional Hill, *Boundary-Layer Meteorology*, 141, 219–243, <https://doi.org/10.1007/s10546-011-9636-y>, 2011.
- 690 Blocken, B.: Computational Fluid Dynamics for urban physics: Importance, scales, possibilities, limitations and ten tips and tricks towards accurate and reliable simulations, *Building and Environment*, 91, 219–245, <https://doi.org/10.1016/j.buildenv.2015.02.015>, 2015.
- Calzolari, G. and Liu, W.: Deep learning to replace, improve, or aid CFD analysis in built environment applications: A review, *Building and Environment*, 206, 108 315, <https://doi.org/10.1016/j.buildenv.2021.108315>, 2021.
- Charbonnier, P., Blanc-Féraud, L., Aubert, G., and Barlaud, M.: Two deterministic half-quadratic regularization algorithms for
 695 computed imaging, in: *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, vol. 2, pp. 168–172, <https://doi.org/10.1109/ICIP.1994.413553>, 1994.
- Davis, N. N., Badger, J., Hahmann, A. N., Hansen, B. O., Mortensen, N. G., Kelly, M., Larsén, X. G., Olsen, B. T., Floors, R., Lizcano, G., Casso, P., Lacave, O., Bosch, A., Bauwens, I., Knight, O. J., van Loon, A. P., Fox, R., Parvanyan, T., Hansen, S. B. K., Heathfield, D., Onninen, M., and Drummond, R.: The Global Wind Atlas: A High-Resolution Dataset of Climatologies and Associated Web-Based
 700 Application, *Bulletin of the American Meteorological Society*, 104, E1507 – E1525, <https://doi.org/10.1175/BAMS-D-21-0075.1>, 2023.
- Doms, G. and Baldauf, M.: A Description of the Nonhydrostatic Regional COSMO-Model. Part I: Dynamics and Numerics, Tech. rep., Deutscher Wetterdienst (DWD), Offenbach, Germany, <https://cosmo-model.org>, 2018.
- Dujardin, J. and Lehning, M.: Wind-Topo: Downscaling near-surface wind fields to high-resolution topography in highly complex terrain with deep learning, *Quarterly Journal of the Royal Meteorological Society*, 148, 1368–1388, <https://doi.org/10.1002/qj.4265>, 2022.
- 705 Dujardin, J., Kahl, A., Krut, B., Bartlett, S., and Lehning, M.: Interplay between photovoltaic, wind energy and storage hydropower in a fully renewable Switzerland, *Energy*, 135, 513–525, <https://doi.org/10.1016/j.energy.2017.06.092>, 2017.
- Dujardin, J., Kahl, A., and Lehning, M.: Synergistic optimization of renewable energy installations through evolution strategy, *Environmental Research Letters*, 16, 064 016, <https://doi.org/10.1088/1748-9326/abfc75>, 2021.
- Eivazi, H., Tahani, M., Schlatter, P., and Vinuesa, R.: Physics-informed neural networks for solving Reynolds-averaged Navier–Stokes
 710 equations, *Physics of Fluids*, 34, 075 117, <https://doi.org/10.1063/5.0095270>, 2022.
- Franke, J., Hellsten, A., Schlünzen, H., and Carissimo, B.: Best practice guideline for the CFD simulation of flows in the urban environment, Tech. Rep. COST Action 732, COST European Cooperation in Science and Technology, <https://hal.science/hal-04181390>, 2007.
- Gasser, J. and Lehning, M.: CFD dataset of wind and pressure fields over Swiss complex terrain with photovoltaic structures and wind turbines (292 OpenFOAM RANS simulations), *EnviDat [data set]*, <https://doi.org/10.16904/envidat.810>, 2026.
- 715 Giral, F., Manzano, Á., Gómez, I., Vinuesa, R., and Le Clainche, S.: GenDA: Generative Data Assimilation on Complex Urban Areas via Classifier-Free Diffusion Guidance, <https://arxiv.org/abs/2601.11440>, 2026.



- Guo, X., Li, W., and Iorio, F.: Convolutional neural networks for steady flow approximation, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 481–490, <https://doi.org/10.1145/2939672.2939738>, 2016.
- Hersbach, H., Bell, B., Berrisford, P., et al.: The ERA5 global reanalysis, *Quarterly Journal of the Royal Meteorological Society*, 146, 1999–2049, <https://doi.org/10.1002/qj.3803>, 2020.
- Jiang, Q., Shu, C., Zhu, L., Yang, L., Liu, Y., and Zhang, Z.: Applications of finite difference-based physics-informed neural networks to steady incompressible isothermal and thermal flows, *International Journal for Numerical Methods in Fluids*, 95, 1565–1590, <https://doi.org/10.1002/fld.5217>, 2023.
- Kahl, A., Dujardin, J., and Lehning, M.: The bright side of PV production in snow-covered mountains, *Proceedings of the National Academy of Sciences*, 116, 1162–1167, <https://doi.org/10.1073/pnas.1720808116>, 2019.
- Kristianti, F., Dujardin, J., Gerber, F., Huwald, H., Hoch, S. W., and Lehning, M.: Combining weather station data and short-term LiDAR deployment to estimate wind energy potential with machine learning: A case study from the Swiss Alps, *Boundary-Layer Meteorology*, 188, 185–208, <https://doi.org/10.1007/s10546-023-00808-y>, 2023.
- Kristianti, F., Gerber, F., González-Herrero, S., Dujardin, J., Huwald, H., Hoch, S. W., and Lehning, M.: Influence of air flow features on alpine wind energy potential, *Frontiers in Energy Research*, 12, 1379 863, <https://doi.org/10.3389/fenrg.2024.1379863>, 2024.
- Kruyt, B., Lehning, M., and Kahl, A.: Potential contributions of wind power to a stable and highly renewable Swiss power supply, *Applied Energy*, 192, 1–11, <https://doi.org/10.1016/j.apenergy.2017.01.085>, 2017.
- Kruyt, B., Dujardin, J., and Lehning, M.: Improvement of wind power assessment in complex terrain: The case of COSMO-1 in the Swiss Alps, *Frontiers in Energy Research*, 6, 102, <https://doi.org/10.3389/fenrg.2018.00102>, 2018.
- Lehning, M., Löwe, H., Ryser, M., and Raderschall, N.: Inhomogeneous precipitation distribution and snow transport in steep terrain, *Water Resources Research*, 44, <https://doi.org/10.1029/2007WR006545>, 2008.
- Lehning, M., Dujardin, J., and Cintas, A.: The value of weather forecasts for optimally placed wind and solar electricity generators, *Frontiers in Energy Research*, 13, 1533 778, <https://doi.org/10.3389/fenrg.2025.1533778>, 2025.
- Li, Z., Zheng, H., Kovachki, N., Jin, D., Chen, H., Liu, B., Azizzadenesheli, K., and Anandkumar, A.: Physics-informed neural operator for learning partial differential equations, *ACM/IMS Journal of Data Science*, 1, 1–27, <https://doi.org/10.1145/3648506>, 2024.
- Lin, C., Tie, R., Yi, S., Liu, D., Zhong, X., Hu, Z., and Li, H.: Reconstructing fine-scale 3D wind fields with terrain-informed machine learning, *Nature Communications*, 17, 3713, <https://doi.org/10.1038/s41467-026-70562-5>, 2026.
- Loshchilov, I. and Hutter, F.: Decoupled Weight Decay Regularization, in: International Conference on Learning Representations (ICLR), <https://api.semanticscholar.org/CorpusID:53592270>, 2019.
- Mott, R. and Lehning, M.: Meteorological modeling of very high-resolution wind fields and snow deposition for mountains, *Journal of Hydrometeorology*, 11, 934–949, <https://doi.org/10.1175/2010JHM1216.1>, 2010.
- Pathak, J., Subramanian, S., Harrington, P., et al.: FourCastNet: A Global Data-driven High-resolution Weather Model using Adaptive Fourier Neural Operators, *arXiv preprint arXiv:2202.11214*, <https://doi.org/10.48550/arXiv.2202.11214>, 2022.
- Raderschall, N., Lehning, M., and Schär, C.: Fine-scale modeling of the boundary layer wind field over steep topography, *Water Resources Research*, 44, <https://doi.org/10.1029/2007WR006544>, 2008.
- Raissi, M., Perdikaris, P., and Karniadakis, G. E.: Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *Journal of Computational Physics*, 378, 686–707, <https://doi.org/10.1016/j.jcp.2018.10.045>, 2019.



- Reynolds, D., Gutmann, E., Kruyt, B., Haugeneder, M., Jonas, T., Gerber, F., Lehning, M., and Mott, R.: The High-resolution Intermediate Complexity Atmospheric Research (HICAR v1.1) model enables fast dynamic downscaling to the hectometer scale, *Geoscientific Model Development*, 16, 5049–5069, <https://doi.org/10.5194/gmd-16-5049-2023>, 2023.
- Rohrhofer, F. M., Posch, S., Gößnitzer, C., and Geiger, B. C.: Data vs. Physics: The Apparent Pareto Front of Physics-Informed Neural Networks, *IEEE Access*, 11, 86 252–86 261, <https://doi.org/10.1109/ACCESS.2023.3302892>, 2023.
- Ronneberger, O., Fischer, P., and Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation, in: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, edited by Navab, N., Hornegger, J., Wells, W. M., and Frangi, A. F., pp. 234–241, Springer International Publishing, Cham, ISBN 978-3-319-24574-4, 2015.
- Rygiel, P., Suk, J., Yeung, K. K., Brune, C., and Wolterink, J. M.: Active Learning for Deep Learning-Based Hemodynamic Parameter Estimation, *arXiv preprint arXiv:2503.03453*, <https://doi.org/10.48550/arXiv.2503.03453>, 2025.
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., Barker, D. M., and Huang, X.-Y.: A Description of the Advanced Research WRF Model Version 4, Tech. Rep. NCAR/TN-556+STR, National Center for Atmospheric Research, Boulder, CO, USA, <https://doi.org/10.5065/1dfh-6p97>, 2019.
- Smith, L. N. and Topin, N.: Super-convergence: very fast training of neural networks using large learning rates, in: *Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications*, vol. 11006 of *Proc. SPIE*, p. 1100612, <https://doi.org/10.1117/12.2520589>, 2019.
- Stathopoulos, T., Zisis, I., and Xypnitou, E.: Local and overall wind pressure and force coefficients for solar panels, *Journal of Wind Engineering and Industrial Aerodynamics*, 125, 195–206, <https://doi.org/10.1016/j.jweia.2013.12.007>, 2014.
- Swiss Federal Office of Energy (SFOE): Wind Atlas of Switzerland, https://www.uvek-gis.admin.ch/BFE/storymaps/EE_Windatlas/, accessed 2026-06-23, 2023.
- Swiss Federal Office of Energy (SFOE): Energy Strategy 2050, <https://www.bfe.admin.ch/bfe/en/home/policy/energy-strategy-2050.html>, accessed June 2026, 2024.
- Tancik, M., Srinivasan, P. P., Mildenhall, B., Fridovich-Keil, S., Raghavan, N., Singhal, U., Ramamoorthi, R., Barron, J. T., and Ng, R.: Fourier Features Let Networks Learn High Frequency Functions in Low Dimensional Domains, *NeurIPS*, 2020.
- Thuerey, N., Weißenow, K., Prantl, L., and Hu, X.: Deep learning methods for Reynolds-averaged Navier–Stokes simulations of airfoil flows, *AIAA Journal*, 58, 25–36, <https://doi.org/10.2514/1.J058291>, 2020.
- Troen, I. and Petersen, E. L.: *European Wind Atlas*, Risø National Laboratory, Roskilde, Denmark, ISBN 87-550-1482-8, 1989.
- van Schaik, B., Reynolds, D., Huwald, H., and Lehning, M.: Resolving three-dimensional wind velocity fields in complex terrain using sequential wind-Doppler LiDAR, CFD and wind turbine measurements – Gotthard Pass, Switzerland, *Open Research Europe*, 5, 9, <https://doi.org/10.12688/openreseurope.19095.3>, 2026.
- Wang, S., Wang, H., and Perdikaris, P.: Learning the solution operator of parametric partial differential equations with physics-informed DeepONets, *Science Advances*, 7, eabi8605, <https://doi.org/10.1126/sciadv.abi8605>, 2021.
- Weller, H. G., Tabor, G., Jasak, H., and Fureby, C.: A tensorial approach to computational continuum mechanics using object-oriented techniques, *Computers in Physics*, 12, 620–631, <https://doi.org/10.1063/1.168744>, 1998.
- Zängl, G., Reinert, D., Rípodas, P., and Baldauf, M.: The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core, *Quarterly Journal of the Royal Meteorological Society*, 141, 563–579, <https://doi.org/10.1002/qj.2378>, 2015.



Zhao, Y., Yang, Y., Zhang, H., and Li, M.: Experimental investigation of wind loads on solar arrays with varying row configurations and tilt angles, *Renewable Energy*, 242, 122 440, <https://doi.org/10.1016/j.renene.2025.122440>, 2025.