Wind Field Reconstruction and Uncertainty Quantification at Wildland Fires Based on Sparse UAV-based Wind Measurements

Mohammad Tavakol Sadrabadi[*], Mauro Sebastián Innocente[*]

Autonomous Vehicles and Artificial Intelligence Laboratory, Coventry University, 7th Floor Friars House, Manor House Drive, Coventry, CV1 2TE, UK

ARTICLE INFO

Keywords:
Wind downscaling
Fire-induced wind
Fire dynamics simulator (FDS)
Physics-constrained Neural Networks
Fluid flow reconstruction

ABSTRACT

Wildfire behaviour is highly influenced by weather, fuel and topography, resulting in highly dynamic propagation patterns. However, a detailed physics-based simulation of this dynamic behaviour can be computationally expensive and time-consuming even on small scales, particularly when accounting for fire-wind interactions. To overcome this limitation, a series of models are developed to provide rapid estimations of fire behaviour by simplifying or ignoring certain physical laws–albeit at the cost of accuracy. These models are often decoupled from the atmosphere to reduce the computational demands, which leads to increased uncertainty in their predictions. Additionally, efforts to improve model reliability by incorporating near-surface wind fields into the model using statistical and dynamical downscaling methods face two challenges including the failure to account for fire-wind interaction and the high computational demand of dynamical methods. Consequently, this study introduces a novel framework that combines UAV-swarm-based wind and temperature measurements with convolutional neural networks (CNN), to estimate the fire-induced near-surface wind field, aiming to capture the fire-wind interaction and its effect on the fire propagation dynamics in a grassland fire without solving the complete set of Navier-Stokes equations. The framework includes a two-step process for wind field estimation including (i) super-resolution reconstruction of the high-altitude wind field from sparse UAV-based measurements, and (ii) high-resolution estimation of the near-surface wind field based on the reconstructed high-altitude wind field. The estimated wind field could then be fed into decoupled wildfire models to replicate the effect of fire-wind interaction on fire propagation. Given the extensive data requirement of deep learning models and lack of access to real-world measured data, this study utilizes synthetic data generated from executing 150 three-dimensional Large Eddy simulations of wildfire propagation in grasslands with varying wind speeds, terrain slopes, vegetation types, and height. The accuracy and uncertainty levels of the trained models are evaluated for different UAV swarm sizes, ranging from 100 to 9 UAVs, as well as various sampling strategies, focusing on the spatial distribution of UAVs above the field. Additionally, the models' reliability are tested under different wind measurement errors by UAV-mounted sensors, varying from 0 to 50%. The obtained results indicate that the developed framework is capable of providing accurate estimations from the near-surface wind field, even under scenarios with a limited number of UAVs, demonstrated through average MAE and RMSE values equal to 0.849 and 1.323 for the U, 0.672 and 1.022 for the V, and 0.551 and 1.01 for the W component of velocity. Uncertainty analysis indicates that even though the average performance of the model remains stable, model uncertainty increases with reducing the size of the swarm. Finally, the investigation of the effect of wind measurement errors on model accuracy and reliability indicates that increased noise levels significantly impact the model's accuracy and uncertainty. However, increasing the swarm size helps to mitigate the effects of measurement noise to a certain extent.

1. Introduction

In recent years, the frequency, scale, and severity of wildfires have undergone a substantial rise globally as a result of the impacts of climate change [1, 2] urging the need for a comprehensive understanding of the wildfire dynamics and its interaction with the atmosphere.

1.1. Fire-atmosphere interaction

Fire-atmosphere interaction refers to the complex interactions between a fire and the surrounding atmospheric environment, which encompasses the interaction between the burning fuels and atmosphere, the energy and mass fluxes between dead and living vegetation, as well as the interaction between the atmosphere and the combustion process as long as the resultant perturbations are larger than the normal ambient variability [3, 4]. During the past century, a large number of studies have documented the influence of the ambient wind or ground-level wind on fire growth dynamics. The first observations of Show[5] suggest a linear fire growth rate over time with a circular perimeter in cases with no ambient wind while reporting an elliptical formation in the presence of wind. [6] noted that the wind contributes to the fire propagation through four procedures (i) facilitating the moisture exchange between the vegetative bed and the atmosphere, (ii) increasing the rate of spread of fire, (iii) carrying ignited embers to the unburnt downstream areas, and (iv) contributing in oxygen flux into combustion front. Tilting the flame, the ambient wind affects the rate of spread of fire (RoS) by changing the distance between the flame and the fuel bed, which enhances both radiative and convective heat transfer [3, 7]. Wildfires in turn enhance the

---

✉ `tavakolsam@uni.coventry.ac.uk` (M. Tavakol Sadrabadi); `Mauro.S.Innocente@coventry.ac.uk` (M.S. Innocente)

🌐 `pureportal.coventry.ac.uk/en/persons/mohammad-tavakol-sadrabadi` (M. Tavakol Sadrabadi); `https://msinnocente.com/` (M.S. Innocente)

ORCID(s): `0000-0002-5938-6310` (M. Tavakol Sadrabadi); `0000-0001-8836-2839` (M.S. Innocente)

surrounding environment by releasing a significant amount of heat as the product of the combustion process. The latent heat released induces buoyant updrafts [8] called fire visible plume as well as the lateral down draft and horizontal vortices surrounding the plume[4]. This upward movement of the air in turn creates a low-pressure area that can cause lateral indrafts from the surrounding atmosphere into the fire region [8]. Hence, a precise definition of turbulent structure and momentum fluxes surrounding the fire is important as they cause spatial and temporal changes in fire spread dynamics [9]. Even low-intensity fires may produce atmospheric turbulence that affects the fire front dynamics [10]. Consequently, this field has gathered considerable attention and a significant number of studies have focused on the characterisation of atmospheric turbulence in the vicinity of fires through large-scale experiments such as the International Crown Fire Modelling Experiment ($ICFME$) [11], field scale experiments [12], and small scale experiments[13]. Another group of studies such as [14, 15, 16] have focused on the numerical simulation to provide a detailed description of the coupled fire atmosphere interaction. Considering that wildfire is a multiscale phenomenon, it should be taken into account that the physics and dynamics of the models are highly reliant on the scales at which the desired events occur. This implies that there are often distinct sets of equations for various scales [17]. A series of models such as FIRETEC [18], and *WFDS* [19], are developed to solve the fluid dynamics equations and thermochemical properties to capture the dynamics of a wildfire. These fully coupled models account for wildfire interaction with the surrounding atmosphere by altering the fire environment through air moisture, temperature, wind speed and direction [17]. Such a detailed solution, although it provides a clear picture of extreme events in wildfires and is very valuable for research purposes, would be inherently computationally expensive, especially for large domains [17], which generally limits the applicability of such models for a majority of applications such as operational fire management tasks. To address the need for decision support tools capable of providing fast and large-scale estimations, a series of faster-than-real-time (FTRT) models such as FARSITE [20], BEHAVE [21], and FireProM-F [22, 23] are developed. These fire propagation models are generally decoupled or partially coupled to the wind to decrease the amount of computation and time required for the solution. Consequently, the level of uncertainty of these uncoupled models increases significantly by increasing the estimation interval [17]. Efforts have been made to improve the accuracy of estimation and reduce the uncertainty of these models through the use of data assimilation [24] and optimisation [25] techniques, or using numerical weather prediction (*NWP*) models to estimate the wind field over the terrain and feeding these estimations into fire propagation models after proper downscaling [26]. However, this method still has some limitations, as these large-scale models may not capture the impact of smaller-scale fires.

1.2. Drones in wildfire management activities

Considering the progression in the capabilities of UAVs during the past years, they have gathered attention from the fire community and have been used in different studies and projects including fire detection, monitoring, and fighting assistance [see 27]. They usually have a series of sensors on board, including the Global Navigation Satellite System (GNSS) and Inertial Navigation System (INS) for *UAV* localization and at least one imagery sensor such as RGB or Infrared cameras [28]. Even though *UAV*s are less suitable for prolonged measurements due to the limitations that contribute to the limited flight time, they have mobility in three dimensions, which makes them flexible and cost-affordable instruments for providing spatiotemporal measurements of the desired variables such as the atmospheric wind [29]. As a result, and due to their advantages over other options available such as wind LIDARs, balloons, and wind measurement towers, an increased number of studies have focused on the use of UAVs for wind field measurement tasks (e.g. [see 30, 31]). Generally, both fixed-wing and rotary-wing UAVs could be used for atmospheric wind measurements, with fixed-wing UAVs more suited to measurement tasks that cover large areas and rotary-wind UAVs mostly suitable for measurements that need hovering on a spot for long periods such as validating wind measurements and in the proximity of the structures [29]. Wind measurements using drones could be performed through the use of external wind sensors or onboard sensors of the avionic system of the *UAV* [32]. The idea of performing coordinated wind measurement with a group of UAVs was presented first by [33] yet the first actual implementation of such a system in the field is performed by [32] that utilizes a group of quadrotors hovering for a long period at certain locations, to provide spatiotemporal wind measurements.

1.3. Wind downscaling/reconstruction

Generally, downscaling refers to the process of obtaining information about physical quantities at a smaller local scale from lower-resolution available data through the use of appropriate refinement methods [34]. Traditionally, wildland fire decision support systems utilise wind data obtained from coarse-scale meteorological service models. Using these data, a domain-average wind field is created and used for fire propagation forecasting [35]. However, some studies have focused on developing and evaluating alternative approaches for this practice in order to support wildfire modelling. [35, 36] compared the effect of three techniques for calculating fine-scale surface winds from *NWP* large-scale models on fire propagation estimations. These methods include(i) using a uniform wind field applied over the whole train, (ii) utilizing 2D mass-conserving models [37, 38], and (iii) utilizing Mass and momentum conserving numerical models for wind field downscaling. Their results indicated that with proper fire model settings, high-resolution input wind fields increase the accuracy of fire estimations. Additionally, it was found that the accuracy of the fire growth simulations increases with the complexity of the wind model.

Downscaling methods can be classified into two broad groups including (i) dynamical and (ii) empirical-statistical approaches [39]. Dynamical downscaling refers to the procedure of using high-resolution mass-conserving or momentum-conserving numerical models to interpret the outputs of *NWP* models on a finer scale. Considering that this method is computationally expensive and timely, especially for large domains, it is mainly restricted to small domains [34]. On the other hand, statistical downscaling efforts to train statistical models between quantities at coarse and fine scales. Such models would have a very short inference time and hence can provide instant estimations when fully trained.

So far, different statistical downscaling methods have been developed [40, 41] yet, with the recent advancements in machine learning (ML) and deep learning (DL) methods, different studies have focused on the use of the state-of-the-art DL algorithms and methods for this purpose [see 42, 43]. In particular, Convolutional Neural networks (CNN) and single image super-resolution (SR) tasks [see 44, 45] have proven to be effective in this manner due to their capabilities in learning multidimensional mappings, especially for spatially distributed data and generation of high-resolution data from low-resolution inputs [34].

This super-resolution procedure is however generally referred to as flow field reconstruction within the context of flow mechanics literature. An example would be to interpret a low-resolution fluid flow image (data matrix) as a collection of sparse sensor measurements. Utilising such an approach, super-resolution analysis could be extended to address the inverse problem of reconstructing global fields from local observations [46]. Different studies have utilised various deep-learning models for flow field reconstruction purposes. Examples include convolutional neural networks mostly with UNet structures (i.e. see [47, 48]), artificial neural networks with autoencoder architectures [49], and physics-informed frameworks [42, 50].

1.4. Uncertainty estimation for flow field reconstruction

Deep learning and machine learning models are frequently used for various types of inference and decision-making procedures. However, assessing the effectiveness and reliability of these AI systems before deploying and utilising them is crucial as these models are susceptible to noise in data (aleatoric uncertainty) and model errors (epistemic uncertainty). Nowadays, uncertainty quantification (UQ) forms the basis of many important decisions; and predictions that lack UQ are typically unreliable [51]. Generally, UQ methods could be categorised into two broad groups including (i) Bayesian methods and (ii) ensemble methods, where Bayesian methods encompass techniques such as the Markov chain Monte-Carlo (MCMC) algorithm, variation inference methods (VI), Monte-Carlo dropout (MCD), etc [51]. However, methods like Bayesian neural networks, MCMC, and the ensemble of models are heavily computationally intensive and time-consuming which limits their utilisation for large datasets and models [51, 52]. As a result, the use of approximation techniques such as MCD [53] and Monte-Carlo batch normalisation (MCBN) [54] which could be used to estimate the epistemic uncertainty of a single trained model has gained traction.

[53] showed that an approximation to the probabilistic deep Gaussian process can be obtained mathematically by using a neural network with arbitrary depth and non-linearities, where dropout is applied before each weight layer. The dropout layer would then be active during model inference where an ensemble of estimations is output to act as the Monte Carlo samples to calculate the mean and variance of the prediction distribution [52, 54]. The Monte Carlo batch normalisation (MCBN) is another technique that could be used to approximate the Bayesian inference in neural networks [54]. In this approach, many mini-batches are created by selecting random samples to go along with a particular query during inference. The predictive distribution is then estimated using the mean and variance of the ensemble of outputs [52].

1.5. Contribution

Reviewing the literature, it could be summarized that up to date, some studies have focused on developing models and approaches to decrease the uncertainty of operational models. One approach involves inputting the near-surface wind field into these models. Various dynamical and statistical methods have been developed to generate the required fine-scale near-surface wind field. Nevertheless, although these methods have been successful in reducing this uncertainty to a certain extent, there are still two issues about these methods that remain unaddressed:

- Utilizing a dynamic downscaling method, although is significantly faster than executing a fully coupled simulation, still requires significant processing capacity, especially for large territories.
- Considering that both statistical and dynamical downscaling methods use the outputs of NWP models or the measurements of synoptic stations as input, the enhancing effect of fire on the surrounding atmospheric flow will not be seen by them, especially for smaller-scale fires.

Considering the above drawbacks, [55] proposed a methodology based on the combination of coordinated real-time *UAV*-swarm-based measurements of the turbulent atmosphere above the fire and the use of deep neural networks for reconstruction and downscaling of the sparsely measured wind field to the high-resolution near-surface wind. Such a method offers the capability of enhancing wildfire model predictions by providing near-surface wind fields approximated from UAV measurements at higher altitudes to the model, effectively mimicking the fire–wind interaction without incurring the hefty cost of solving the Navier-Stokes equations [55]. Readers are kindly referred to the original paper for a detailed description of the results of improving fire propagation simulations.

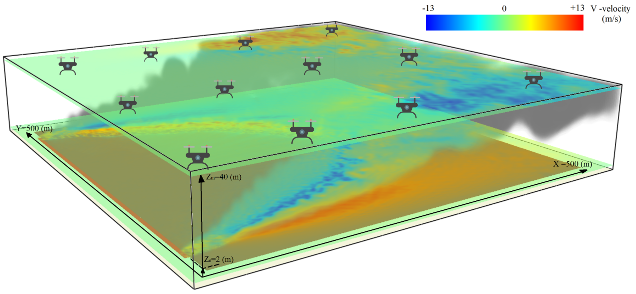


Figure 1: Schematic view of the fire propagation and distribution of the transverse component of velocity (V) at 2m *AGL* and 40m *AGL* along with the *UAV* swarm .

This paper, however, focuses on the detailed development of the wind reconstruction and estimation procedure conceptually proposed in [55] including details on 3D simulations carried out to provide synthetic data for training data-driven models, model architecture development and assessment. Finally, an uncertainty quantification study is performed utilising Monte-Carlo batch normalisation (MCBN) and Monte-Carlo dropout (MCD) techniques to perform an approximate Bayesian inference over the neural network parameters, facilitating the estimation of epistemic uncertainty of predicted wind fields.

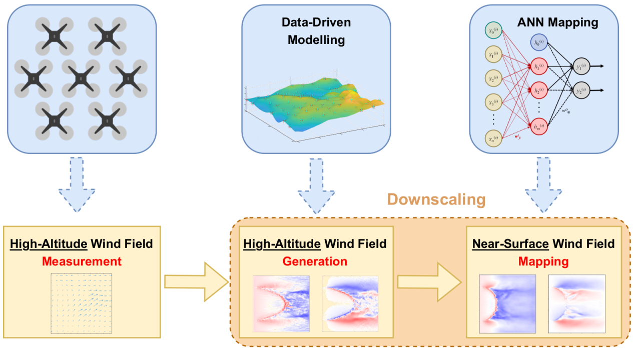


Figure 2: Schematic description of the proposed methodology adopted and modified from [55]

2. System description

This paper develops a data-driven system for downscaling wind patterns in the vicinity and above grassland fires to improve the accuracy and reduce the uncertainty of fire propagation estimations based on real-time measurements carried out by a swarm of UAVs. A schematic view of the system is presented in Fig 1. Given that the fire plume in a wildland fire is tilted forward due to the ambient wind velocity, the turbulent fields near the surface and at the measurement height differ significantly. Therefore, downscaling atmospheric turbulence in the presence of fire cannot be directly addressed through a single-image super-resolution task. Hence, the developed system consists of three steps to provide real-time wind field maps near the terrain surface. The steps include (i) real-time UAV-swarm-based measurement of the turbulent wind above the fire, (ii) reconstruction of a high-resolution wind field at UAV flight altitude from sparse point measurements, and (iii) estimation of the high-resolution wind field close to the terrain surface based on the reconstructed wind field at UAV flight altitude. It is important to note that while the downscaling process is typically equated with the super-resolution or reconstruction step in the literature, this study considers downscaling to encompass both the super-resolution/reconstruction step and the estimation of the high-resolution near-surface wind field from the reconstructed high-altitude wind field. A schematic description of the proposed methodology is presented in Figure 2. Note that this study addresses both steps in downscaling blocks utilising convolutional neural networks.

2.1. Swarm-based wind measurements

The first component of the system is the *UAV* swarm which is responsible for the measurement of wind at the determined measurement height (Z_m) which is assumed to be equal to 40m above ground level (*AGL*) in this study. The *UAV* swarm is also responsible for determining the location and the temperature of the firefront. The swarm may consist of a group of rotary-wing or fixed-wing small aircraft, allowing for variability and flexibility in measurement strategies based on the specifications and limitations of the UAVs used. The most common sensors used in UAVs include pitot tubes and multi-hole pressure probes, mechanical and sonic anemometers, and LiDAR systems [55] where the Wind speed measurement errors carried out by quadrotors are reported to include errors in the range of 20–50% [56]. Irrespective of the source of measurements, the general limitation in wind measurement is that the Nyquist sampling theorem should be fulfilled to prevent aliasing problems. This is more problematic, especially when using fixed-wind *UAVs* which normally fly at higher speeds (e.g. see [57]). Besides, the turbulence induced by the fire is usually comparable to or higher than the ambient atmosphere which requires a higher sampling rate compared to the normal atmospheric measurements to achieve the desired resolution for turbulent fluctuations. Even though determining such sampling rates is out of the context of this study, implementing methods such as the Nyquist-based adaptive sampling [58] could be beneficial in this manner.

2.2. Wind downscaling

As mentioned, this study assumes that wind downscaling encompasses both high-altitude wind field reconstruction and estimation of the near-surface wind field. Utilising the sparsely measured wind values at each time step as input data into the model, a high-resolution map of the wind field at the measurement altitude (Z_m=40 m) is generated which is in turn used to estimate the near-surface wind field Z_e=2 m *AGL*. Considering that a high-res wind field is generated for the Z_m at the previous step through a super-resolution method, the equivalent high-res wind field for Z_e could be generated through a regression procedure. To achieve this, the study employs an image-to-image translation technique,

Table 1
Overview of the simulated Geometric and physical properties of the vegetation, terrain, and wind

Veg model	Veg height (m)	Terrain Slope (degrees)	U_{10} (m/s)	N. simulations
Veg_1	0.2, 0,5, 10	-40%(-21.8) , -20%(-11.31) , 0 (0), +20%(-11.31), +40%(-21.8)	4, 6, 8, 10, 12	75
Veg_2	0.2, 0,5, 10	-40%(-21.8) , -20%(-11.31) , 0 (0), +20%(-11.31), +40%(-21.8)	4, 6, 8, 10, 12	75

Table 2
Summary of fuel physical properties and thermal decomposition coefficients

Property	Unit	Veg_1(Veg_2)	Reference
Area to Volume Ratio (σ)	m^{-1}	9,770	[59]
Bulk density (ρ_b)	$kg.m^{-3}$	1.313	[59]
Fuel	-	Cellulose	[16]
Fuel Density (ρ)	$kg.m^{-3}$	512	[16]
Moisture content (M)	%	6.3	[59]
Specific Heat	$kJ\ kg^{-1}\ K^{-1}$	2.1	[60]
Conductivity	$kJ\ kg^{-1}\ K^{-1}$	0.1	[16]
Heat of Evaporation (H_{H_2O})	$kJ\ kg^{-1}$	2259	[61]
Heat of Combustion (H_c)	$kJ\ kg^{-1}$	17,400	[62]
Heat of Pyrolysis (H_{pyr})	$kJ\ kg^{-1}$	418	[61]
A_{pyr}	s^{-1}	1040 (24550)	[63] ([64])
E_{pyr}	$J.mol^{-1}$	61041 (58200)	[63] ([64])
Char Yield (ν_{char})	$kg\ kg^{-1}$	0.25 (0.23)	[62]
A_{char}	$kg.m^{-2}.s^{-1}$	465	[65]
E_{char}	$J.mol^{-1}$	68000	[65]
Ash Yield (ν_{ash})	$kg\ kg^{-1}$	0.04	[62]
Obukhov Length (L)	m	−500	[16]
Roughness Length(z_0)	m	0.03	[16]
Drag Coefficient (c_d)	-	2.8	[62]
Soil Specific Heat	$kJ\ kg^{-1}\ K^{-1}$	2.0	[16]
Soil Conductivity	$W\ m^{-1}\ K^{-1}$	0.25	[16]
Soil Density	$kg\ m^{-3}$	1,300	[16]
Relative Humidity	%	40	[16]

which generates the required wind field at 2 m *AGL* from the wind field reconstructed from sparse measurements at 40 m *AGL*.

3. Methods and dataset

Considering that actual measurements of wind velocity in wildfire situations are challenging to obtain and deep learning models require extensive amounts of data for training procedures, this study employs coupled three-dimensional computational fluid dynamics (CFD) simulations of wildland fire propagation to act as synthetic data. Hence, a total number of 150 three-dimensional simulations are performed in order to provide sufficient data required to train the models. All simulations are performed utilising the Fire Dynamics Simulator (FDS) a solver for the Navier-Stokes equations appropriate for low-speed, thermally-driven flows developed by the National Institute of standard and technology (NIST)[62].

3.1. Simulation parameters and boundary conditions

An overview of the physical and geometrical parameters of the simulation conditions including wind velocity, vegetation height and type, and terrain slope are presented in Table 1. Each vegetation type is utilised in 75 field-scale simulations carried out with three different vegetation heights of H_g, five different terrain slopes (S), and five different ambient wind speeds (u_{10}).

The fuel is modelled utilising the lagrangian particle model (LPM) within the FDS which represents the fuel as a group of Lagrangian particles heated by convection-radiation heat transfer [62]. The details of the physical and thermal parameters of fuel utilised in this study are presented in Table 2.

The ambient wind is introduced into the model utilising the Monin-Obukhov similarity theory where the wind speed (u_z) and potential temperature change with height as below:

$$u(z) = \frac{u_*}{k}\left[ln\left(\frac{z}{z_o}\right) - \Psi_m\left(\frac{z}{L}\right)\right] \tag{1}$$

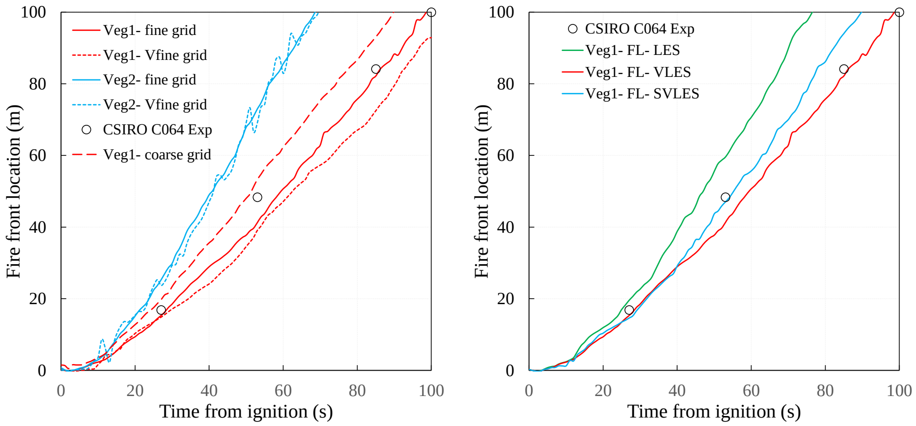
$$\theta(z) = \theta_0 + \frac{\theta_*}{k}\left[ln\left(\frac{z}{z_o}\right) - \Psi_h\left(\frac{z}{L}\right)\right] \tag{2}$$

Where u_* is the friction velocity, k is the von Karman constant equal to 0.41, z_0 is the aerodynamic roughness length, θ_* the scaling potential temperature, θ_0 ground level temperature, L is the Obukhov length, and Ψ_h and Ψ_m represent similarity functions. An unstable stratified condition is assumed to represent the conditions of a warm day aligned with the experimental conditions of the reference experiment.

The turbulent flow field within the domain is simulated utilising the very large eddy simulation (VLES) model with Deardorff's sub-grid scale (SGS) model for turbulent eddy viscosity closure terms. To replicate the turbulent nature of the natural atmospheric winds, this study utilises the synthetic eddy method (SEM)[66] that introduces random eddies into the domain as adopted in relevant studies (i.e. see [67]). Considering that an accurate definition of the eddy characteristics requires measurements of the turbulent Reynolds stress within the ambient wind and canopy height which are not applicable here, an arbitrary value of 10% turbulence intensity is determined here similar to [67]. Utilised boundary Conditions include a no-slip boundary condition for the ground surface and 'Open' boundary conditions for the rest of the boundaries.

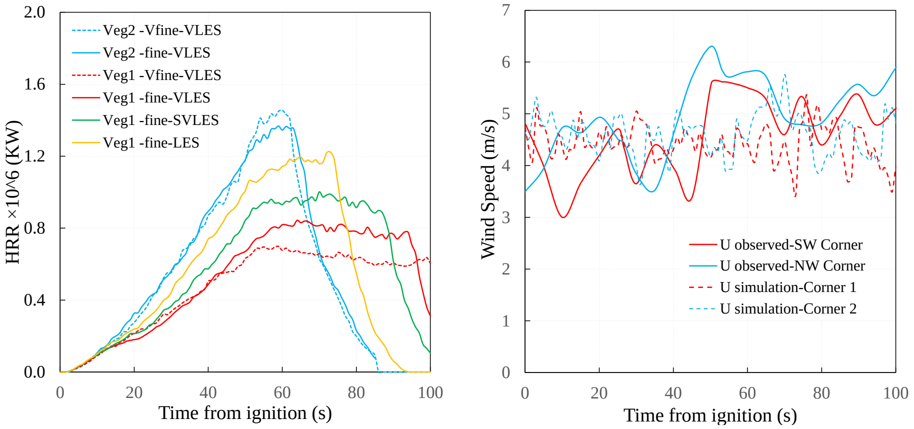
3.2. Grid sensitivity and model validation

Numerical simulations are validated against the field scale wildfire propagation experiments carried out by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) during July and August 1986 with constant high daily temperatures. The reference experiment here is the C064 experiment, carried out in a 100×100 m field covered with Kerosene grass. Experimental conditions include



(a) Effect of grid size on Fire front location compared to experimental data

(b) Effect of turbulence model on Fire front location



(c) Effect of grid size and turbulence model on obtained HRR

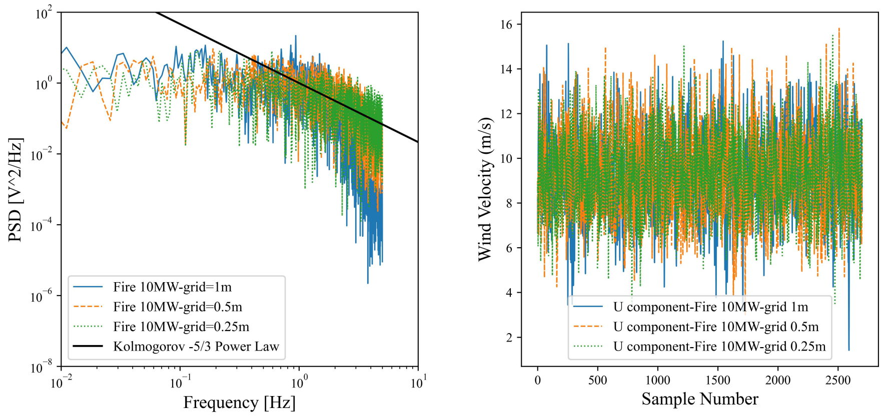
(d) Comparison of measured and simulated wind velocity (U component) during the experiment

Figure 3: Model sensitivity and validation analysis

wind speed u_2=4.6 m/s, temperature T=32°C, and H_g=0.21 m, which resulted in a rate of spread of fire (RoS) equal to 1.2 m/s [59]. In C064 experiments, the fire was ignited using two field workers starting from the centre of the upwind boundary and walking in the opposite direction of each other which is modelled utilizing a strip of burnable grass with a specified rate of spread along the strip.

The sensitivity of the estimated RoS to the grid size is studied by monitoring the firefront location at different time steps after the fire ignition at simulation scenarios comparable to the experimental conditions as depicted in Figure 3a. Three different grid sizes of $1 \times 1 \times 1$ m^3 (coarse), 0.5×0.5×0.5 m^3 (fine), and 0.25×0.25×0.25 m^3 (very fine) were considered for domain discretisation. Furthermore, the effect of the turbulence model on the provided fire front location and heat release ratio (HRR) are also studied through comparing the model estimations while utilising different turbulence models including Simple Very Large Eddy Simulation (*SVLES*), Very Large Eddy Simulation (*VLES*), and Large Eddy Simulation (*LES*) as depicted in Figure 3b and 3c. Finally, Figure 3d provides a comparison between the simulated time series of wind (U component) with the measured values during the experiment. The provided results indicate a general agreement between the simulation and the measured values though a significant shift in experimental values is observed during the experiment.

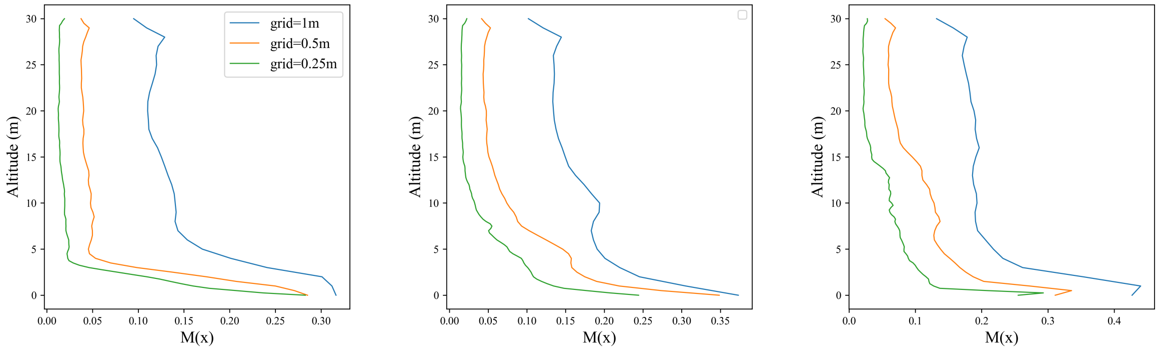
Even though the estimations of the fire-front location agree with the experimental measurements while utilising the chosen grid size of 0.5 m, it is not guaranteed that this grid size is fine enough to adequately resolve a high fraction of the turbulent kinetic energy and ensure a valid Large Eddy Simulation (LES). Consequently, a posteriori grid quality analysis is conducted. This includes studying the effect of grid size on turbulence characteristics by calculating the spectral density of turbulence as well as evaluating the measure of turbulence resolution $M(\mathrm{x})$. The power spectral density is an important parameter, aiding in characterising the distribution of kinetic energy across different frequencies or scales of turbulent motion. Figure 4 presents the effect of grid size on the estimated spectral density (4a) and wind velocity (u_2) time series (4b) measured at x=9 m downstream of a 10MW/m pool fire with a width of 1 m at a



(a) Effect of grid size on PSD distribution

(b) Effect of grid size on wind velocity time-series

Figure 4: Effect of grid size and Fire Intensity on Spectral Density of turbulence



(a) x=0 (fire location)

(b) x=10 m downstream of fire

(c) x=20 m downstream of fire

Figure 5: Effect of grid size on the measure of turbulence resolution $M(\mathrm{x})$ downstream of a 10 MW fire

sampling rate of 10HZ for 260 seconds, 2 meters above the ground level. It should be mentioned that the measurement point falls within the fire plume and hence reflects the fire-induced wind velocities. Generally, it could be observed that the simulated wind time series and the power spectral density distribution estimated using a 0.5 m grid and 0.25m grid are comparable, indicating that the 0.5 m grid might be sufficient for our purpose. It should be highlighted that the observed fire intensities in our experiments span a wide range yet, the 10MW fire here is selected as an average representative of obtained fire conditions during simulations for the calculation of PSD and $M(\mathrm{x})$.

The measure of turbulence resolution is a scalar quantity which could be calculated as follows [62]:

$$M(x) = \frac{k_{\text{sgs}}}{k_{\text{sgs}} + \text{TKE}} \tag{3}$$

where TKE is the turbulent kinetic energy and could be calculated as follows:

$$\text{TKE} = \frac{1}{2}\left((\overline{u} - u')^2 + (\overline{v} - v')^2 + (\overline{w} - w')^2\right) \tag{4}$$

and K_{sgs} is the subgrid kinetic energy and is estimated from Deardorff's eddy viscosity model. It is also suggested that a valid Large Eddy Simulation is obtained such that $M(\mathrm{x})<0.2$ [62].

Figure 5 presents the effect of grid size on the calculated values of the $M(\mathrm{x})$ at different locations downstream of the 10MW fire. It could be concluded that even though the best results are obtained when using the e0.25 m grid size, at all distances downstream of the fire, an approximate amount of 80-85% of turbulent kinetic energy is resolved by the 0.5 m grid except very close to the surface. This hence is a strong indicator that the 0.5 m grid is fine enough for a valid large eddy simulation of wildland fire propagations.

Consequently, it could be observed that a combination of a grid size of 0.5 m and the VLES turbulence model outputs satisfactory results and hence are utilised for the main simulations. It is however acknowledged that the results are grid sensitive to a certain level and a true grid convergence has not been achieved, yet, practical considerations including the large span of the vegetated area, as well as the need for

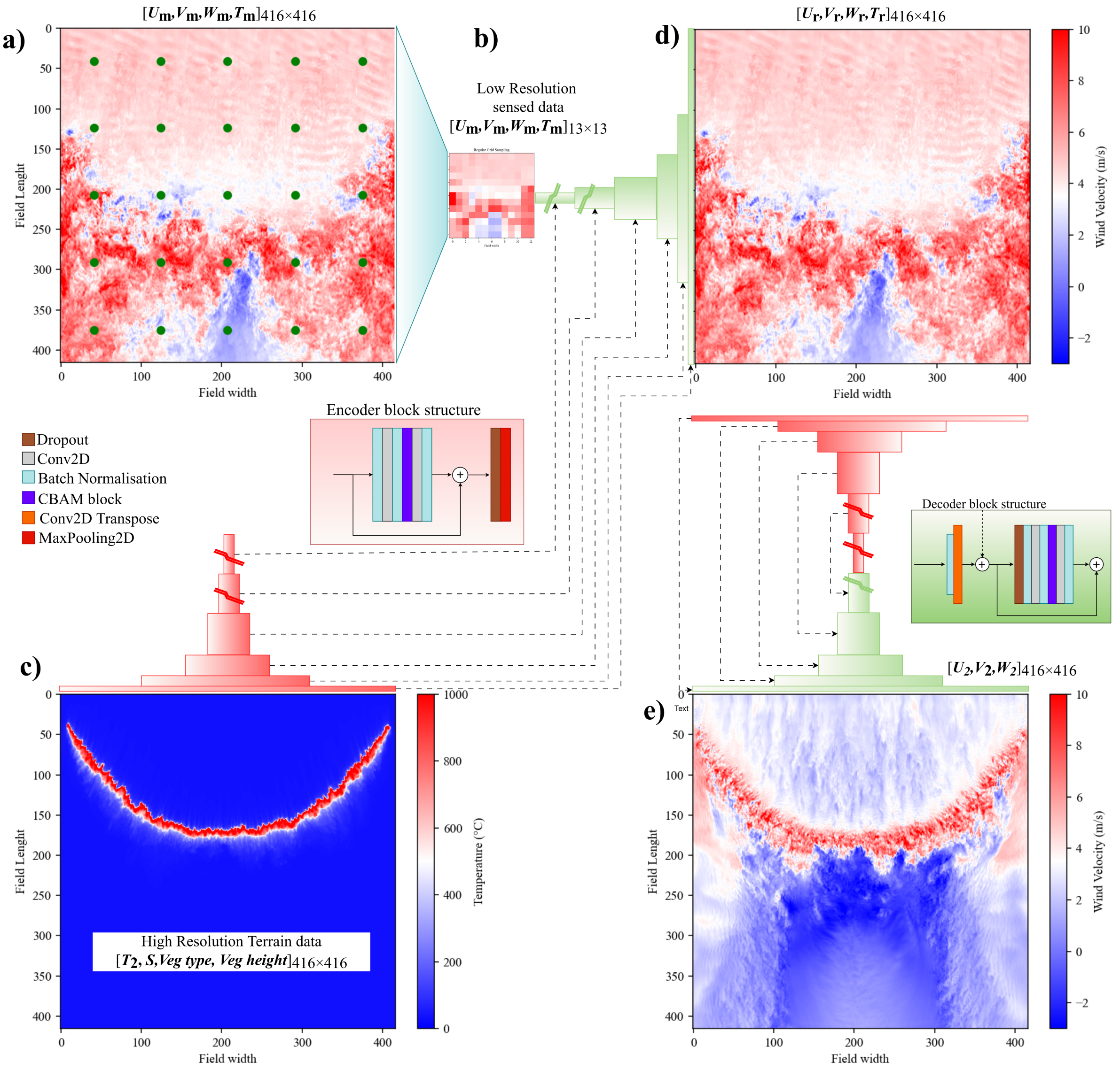


Figure 6: Architecture used for reconstructing the high-resolution wind field from low-resolution sensor data, showing: (a) high-altitude wind field sampled by UAVs in a structured formation, (b) generated low-resolution wind field as input to the super-resolution (SR) model, (c) terrain and fire data used as additional information for wind field reconstruction, (d) reconstructed high-altitude wind field, and (e) estimated near-surface wind field

a considerable height of domain, prevent the utilisation of finer grid size. Thus, the final simulation domain consists of a rectangular block of 600× 320×60 m in x, y and z directions, respectively. A burnable area of 200×200 m is considered in this study to provide sufficient space for the formation of a quasi-steady state, especially for upslope and high wind conditions. Hence, to optimise the computational cost, the domain is discredited utilising a non-uniform grid with grid sizes of $0.5 \times 0.5 \times 0.5$ m^3 for the vegetated area and its adjacent areas (±4 m) up to 44 m high. Upstream and downstream areas up to a distance of ±30 m are discredited with $1 \times 1 \times 1$ m^3 where the rest of the domain is discredited utilising $2{\times}2{\times}1$ m^3 cells. Thus, for each model, the domain is discredited, utilising a total number of 16687680 grid cells parallelised on 32 CPU cores for the duration of the simulation.

3.2.1. *Model architecture and training procedure*

Architecture Briefly, this study effort to estimate the spatial distribution of the near-surface (2m *AGL*) wind velocity components of u_2, v_2 and w_2 in the presence of an active grassland fire from the sparsely measured values (u_{40}, v_{40} and w_{40}) utilising a swarm of UAVs. As previously described, the problem is addressed through supervised training of two convolutional neural networks. Figure 6 presents the overall workflow of the models and data utilised in model training. The first model is trained to reconstruct the high-resolution wind and temperature field at 40m *AGL* from the sparsely measured or low-resolution input data. The low-resolution input data are sampled from the original high-resolution simulations so that each sampled pixel represents a point measurement carried out by an individual UAV. Three different numbers of UAVs (n) are considered in this study so that $n \in (25, 49, 100)$. Sampled pixels $P \in \mathbb{R}^{n\times4}$ are then used to generate the low-resolution input file $I_{\text{lr}} = \mathcal{F}_{\text{interp, KNN}}(P) \in \mathbb{R}^{13\times13\times4}$ through linear interpolation for areas that fall between at least two measurements and assigning values utilising the nearest available value (KNN) to outer regions (figure 6a-b). Extra information including the terrain slope (S) in degrees, the surface temperature ($T_{surf=2}$), Vegetation type (G_t) and the vegetation height (H_g) at corresponding coordinates of the terrain are inputted in the model through an encoder branch and is concatenated with the corresponding upsampled decoder block of the SR model (figure 6c&d). Hence, the output of the super-resolution model would be $I_{\text{sr}} = \mathcal{F}(I_{\text{lr}}, I_{\text{terrain}})$ where $I_{sr} = (u_{40}, v_{40}, w_{40}, T_{40}) \in \mathbb{R}^{416\times416\times4}$, and $I_{terrain}$=(T_2, S, H_g, G_t).

The second model is trained on high-resolution input-output pairs of flow field at 40m AGL and the corresponding values at 2m AGL so that to estimate the near-surface wind field $I_{\text{surf}} = \mathcal{G}(C)$ where $I_{\text{surf}} = (u_2, v_2, w_2)$ from a concatenation of reconstructed high altitude wind field and the high-resolution data corresponding to the surface temperature, vegetation height and type, and terrain slope $C = \text{concat}(I_{sr}, I_{\text{terrain}}) \in \mathbb{R}^{416\times416\times8}$ (see figure 6d-e).

The utilized structure for both networks resembles a UNet with Residual blocks [68] in both encoder and decoder branches empowered with the Convolutional Block Attention Module (*CBAM*)[69] blocks. A sensitivity analysis is performed and the optimal number of model parameters is selected based on a trade-off between obtained accuracy and training time. The activation function for the convolutional layers includes $tanh$ activation for all layers except the output layer which uses the $linear$ activation function. It should be emphasized that the linear activation showed better performance compared to $LeakyReLU$. All convolutional layers use zero padding, a $he-normal$ kernel initializer, and a kernel size of 3×3 except for the first convolutional layer after the input layer, which benefits from a 7×7 kernel, and the output layer which uses a 1×1 kernel. Figure 6 also presents the structure of the encoder and decoder blocks in both models. All blocks include a dropout layer with a dropout ratio of 5% which is meant to assist in improving the model generalisability and uncertainty quantification analysis. All Models are then trained for 200 epochs with MAE loss function and the Adam optimizer with an initial learning rate of 0.0006 and an exponential decay rate of $beta_1$=0.9. During the training, accuracy metrics are observed after each epoch and the best models are saved for later use. To evaluate the performance of neural networks, the mean absolute error MAE, and root mean squared error $RMSE$ values are calculated as follows:

$$RMSE(y, y') = \sqrt{\frac{1}{n}(\sum_{i}^{n} |y_i - y'_i|^2)} \tag{5}$$

$$MAE(y, y') = \frac{1}{n}(\sum_{i}^{n} |y_i - y'_i|) \tag{6}$$

Where y and y' are the ground truth and estimated values respectively.

Concerning probabilistic performance, three additional metrics including the standard deviation (Std), the Prediction Interval Coverage Probability (PICP), and the Prediction Interval Normalized Average Width (PINAW) are utilised to quantitatively assess the model performance. The standard deviation is a measure of variation or dispersion in a set of data values and can be calculated as follows:

$$Std = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N-1}} \tag{7}$$

where x_i is the value of $i_t h$ data point, μ is the population mean, and N is the number of datapoints.

The PICP on the other hand, is a fundamental characteristic of prediction intervals (PIs) and is calculated as the fraction of target values covered by the estimated PIs[70]:

$$PICP = \frac{1}{N}\sum_{i=1}^{N} \epsilon_i; \text{and } \epsilon_i = \begin{cases} 1 & \text{, if } \hat{k}_i \in [L_i, U_i] \\ 0 & \text{, otherwise} \end{cases} \tag{8}$$

where L_i and U_i are the estimated lower and upper limits at the i_{th} array, and N is the number of arrays. Generally, a larger PICP indicates more targets covered by the predicted prediction intervals (PIs). However, it does not necessarily denote an accurate model. Apart from the coverage probability, the width of the estimated PIs is an important metric, defining the quality of the PIs, so that for equal PICPs, a smaller PINAW would normally imply a better model. PINAW measures the average width of the prediction intervals normalized by the range of the actual observations and can be calculated as follows [70]:

$$PINAW = \frac{1}{N}\sum_{i=1}^{N} \frac{U_i - L_i}{Y_{\max} - Y_{\min}} \tag{9}$$

where $Y_{\max}$ and $Y_{\max}$ are the maximum and minimum values of the actual observations, respectively.

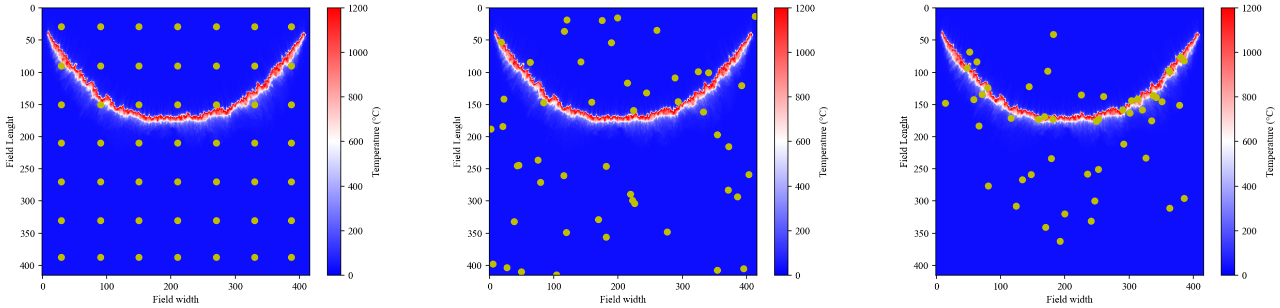
3.3. Dataset

As described before, a total number of 150 simulations are carried out with different combinations of wind speeds, terrain slopes and vegetation height and types. The wind field at 40m and 2m *AGL* (parallel to the terrain) along with the temperature field at the surface are extracted at different time steps after the start of ignition and are used for training the network. The dataset includes 24568 pairs of input-target

Table 3
Effect of the number of model parameters on the SR performance for n=100, utilising first sampling strategy

	model	N.params	U* (m/s)		V* (m/s)		W* (m/s)		T* °C	
			MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Attention	1	1,030,358	0.744	1.277	0.724	1.241	0.875	1.535	3.860	14.891
	2	3,901,084	0.735	1.268	0.717	1.225	0.867	1.514	3.841	14.632
	3	7,270,584	0.732	1.268	0.715	1.223	0.866	1.525	3.889	15.142
	4	15,334,136	0.736	1.274	0.720	1.234	0.864	1.510	3.797	14.626
No-Attention	5	3,901,084	0.741	1.283	0.719	1.230	0.869	1.523	3.857	14.814
Att+weak PI	6	3,901,084	0.737	1.275	0.721	1.235	0.868	1.516	3.845	14.683

* U(m/s) ∈ [-18.7, 55.3], V (m/s)∈ [-38.4, 37.7], W (m/s)∈ [-22.2, 52.1]



(a) Regular grid sampling (b) Random sampling (c) Probability-based random sampling

Figure 7: Wind sampling strategies

data of which 17343 are randomly adopted for training, 3652 for validation, and 3573 for the test.

As mentioned before, the low-resolution input data to the SR model $I_{\text{lr}} \in \mathbb{R}^{13\times13\times4}$ are generated from the corresponding high-resolution images by sampling a prescribed number of pixels (25,49,100) and then creating a continuous response surface by performing linear interpolation between measured pixels and assigning nearest neighbour value to extrapolation areas to form a structured 13×13 low-resolution data matrix for each component of flow $(u_{40}, v_{40}, w_{40}, T_{40})$. To study the effect of sampling distribution on the system performance, three different sampling strategies are adopted in this study to provide the low-resolution input data as depicted in Figure 7. The first strategy as shown in Figure 7a assumes that the UAVs are positioned on a structured grid of points with equal distances and provide a regular grid of sampled pixels. Note that the yellow points reflect the fixed position of UAVs for the whole duration of the simulation. Two other strategies assume that measurements are carried out by UAVs while flying above the domain and not positioned in a structured manner. Hence, the second strategy as shown in Figure 7b is a purely random sampling of the 416×416 flow field. In contrast, the third strategy assumes that pixels with higher near-surface temperatures have a higher probability of being sampled and is defined as follows:

$$P(x, y) = \frac{|T^{0.7}(x, y)|}{\sum_{(i,j)} T(i, j)} \tag{10}$$

where $T(x, y)$ is the temperature at pixel (x, y) and P(x,y) defines the probability of that pixel being sampled. It should be highlighted that the value of 0.7 is selected arbitrarily to provide a visual balance between points sampled over the fire line and samples from the rest of the domain. This third strategy simulates a scenario where a greater number of UAVs are concentrated over the fire line to perform firefighting operations, while other UAVs are either en route to the fire line to discharge their payloads or returning to refill their payloads, battery, etc.

For the first strategy, only one low-resolution image is generated per each high-resolution image. Hence, models are trained on 17343 pairs of data, validated on 3652 pairs and tested on 3573 data pairs. In contrast, the second and third strategies produce varying distributions of sampling points for each image, leading to different low-resolution images every time the sampling process is repeated. This variability reflects the numerous possible distributions of UAVs over the terrain at any specific time, which can have infinite configurations. Therefore, for the second and third strategies, a sensitivity study is conducted to examine the effect of the number of low-resolution samples generated per each high-resolution flow field on the model's generalizability and accuracy in reconstructing a randomly sampled flow field. This is however discussed in more detail in the next section.

4. Results and Discussion

The models are expected to learn the highly complex behaviour of the flow and the relationship between the high-altitude and near-surface flow fields, which is inherently intricate. First, the shape of the fire front changes over time, affecting the turbulent structure of the wind field at both measurement height (Z_m) and near the surface (Z_e) implying a highly dynamic environment. Additionally, by comparing the near-surface values to those from 40m AGL, it can be concluded that the wind fields are inherently different at measurement and estimation heights due to the ambient wind tilting the fire plume forward and creating complex behaviour at the plume flanks. The highest velocity regions near the surface are located immediately after the flaming front while the max velocity at (Z_m) is observed at a significant distance downstream of the fire location. Furthermore, the velocity field near the terrain surface includes two counter-rotating wake vortices downstream of the fire increasing the complexity of the flow field.

The models are designed to be time-independent as estimating the previous or future state of the wind is not of interest to this study. Meanwhile, it is intended to be capable of estimating the distribution of the near-surface wind velocity components in real time from sparse measurements carried out by the UAV swarm. To prevent confusion, this section is divided into two parts where the first part discusses the experiments regarding the super-resolution model, and the second part explains the results obtained from a second model trained to estimate the high-resolution near-surface wind field.

4.1. High altitude flow field reconstruction

This section presents the accuracy metrics and uncertainty measures obtained from the models trained for high-altitude wind field reconstruction.

4.1.1. Model sensitivity analysis

A sensitivity analysis was carried out to evaluate the effect of various parameters on the accuracy of super-resolution output. The number of filters in each layer of the network is the first parameter examined, focusing on how the quantity of the model's trainable parameters influences SR output accuracy. Table 3 presents the results for this preliminary examination. It is indicated that increasing the number of trainable parameters from 1M to 4M has impacted the model accuracy positively while further increase in model depth has resulted in a decline in the model performance. Comparing Model 2 and Model 5 highlights the impact of utilising Convolutional Block Attention Module (CBAM) blocks which has resulted in improved accuracy of the model outputs. To include some physics within the model training procedure, a weakly physics-informed (PI) version of the model is trained by incorporating the incompressible version of the continuity equation into the loss functions so that:

$$Loss_{\text{total}} = Loss_{\text{residual}} + |\nabla \cdot \vec{V}|$$

Table 4
Effect of the number of UAVs on the SR performance utilising first sampling strategy

N. UAV	U* (m/s)		V* (m/s)		W* (m/s)		T* C	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
25 (5*5)	0.767	1.341	0.745	1.285	0.896	1.589	4.295	16.278
49 (7*7)	0.755	1.315	0.735	1.265	0.882	1.558	4.091	15.613
100 (10*10)	0.735	1.268	0.717	1.225	0.867	1.514	3.841	14.632

* U(m/s) ∈ [-18.7, 55.3], V (m/s)∈ [-38.4, 37.7], W (m/s)∈ [-22.2, 52.1]

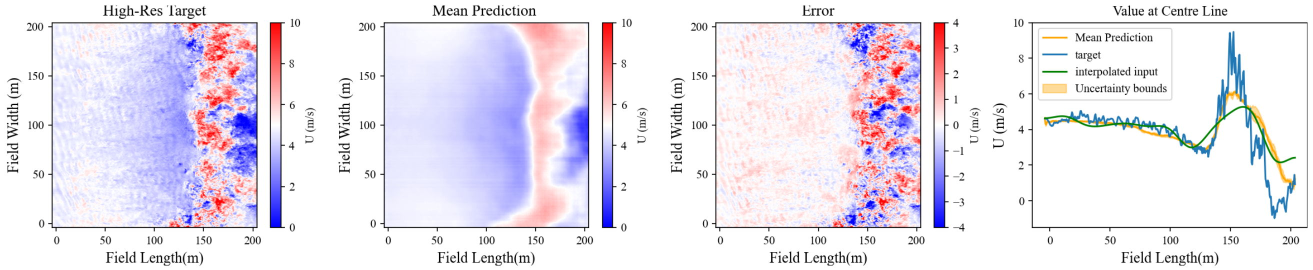
Where $Loss_{\text{total}}$ is the weak PI loss utilised to train the model. However, utilising this approach showed no positive impact on model performance as indicated in table 3 despite increasing the training time and complexity. Consequently, the second model architecture with 4M trainable parameters with 16,32,64,128,256 filters in convolutional layers within the first to fifth encoder and decoder blocks of UNet architecture, is adopted as the optimal architecture for the rest of the examinations.

The second examination involves evaluating the effect of the number of UAVs (sampling points) on the accuracy of the super-resolution output. Consequently, as previously discussed, three different numbers of UAVs including 25, 49, and 100 were considered while generating low-resolution inputs and one model was trained for each UAV quantity. Table 4 presents the effect of the number of UAVs or sampling points while generating the low-resolution input utilising the first sampling strategy, on the output accuracy of the super-resolution model. It can be observed that increasing the number of UAVs results in a marginal improved super-resolution accuracy. While this is not a surprising outcome, it should be highlighted that the difference between the estimated MAE values remained within 4% for velocity components and 10% for temperature, while the number of UAVs changed from 25 to 100, a 4x increase.

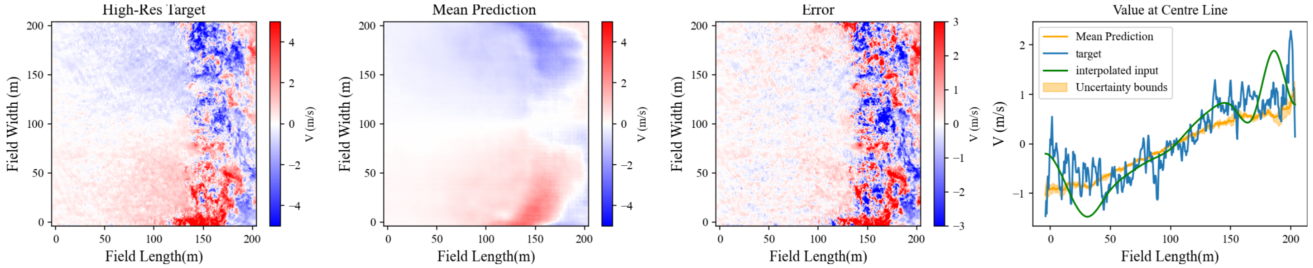
The final examination investigates how training models with various low-resolution versions of a single high-resolution flow field affect model performance. Hence, three different models were trained utilising varying numbers of low-resolution input samples, ranging from 2 to 5, for each high-resolution wind field as presented in Table 5. The models are trained to treat each low-resolution and high-resolution pair as a distinct input pair. For example, the training time for a model with 5 input samples is five times that of a model with a single input sample. Consequently, it could be concluded that while increasing the number of low-resolution inputs slightly increases model performance and generalizability, it also increases the model training time linearly. This hence would make it impractical to account for a high number of inputs while training the primary model on the whole dataset.

4.1.2. Model performance and epistemic uncertainty

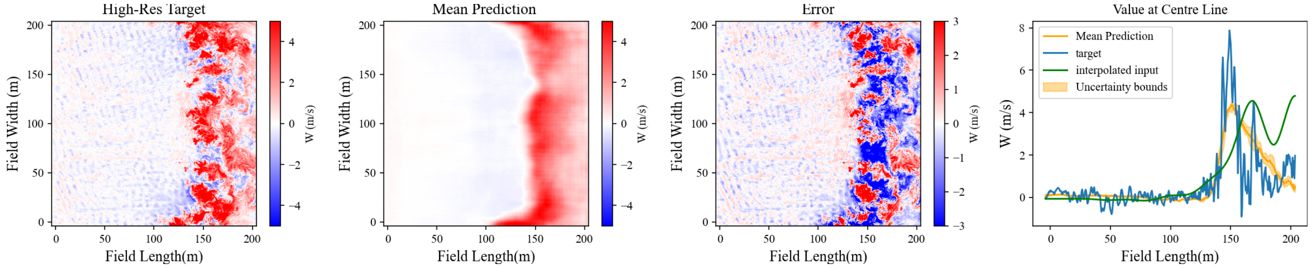
The main model, as previously mentioned, is trained to ensure accurate super-resolution estimations regardless of the sampling strategy or UAV number and distribution. Hence, the training dataset consists of low-resolution and



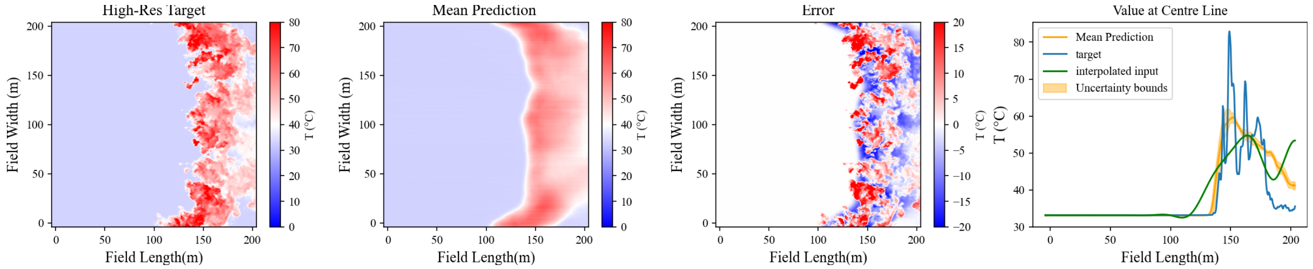
(a) Reconstruction results for U (m/s) component of velocity



(b) Reconstruction results for V (m/s) component of velocity



(c) Reconstruction results for W (m/s) component of velocity



(d) Reconstruction results for temperature T °C

Figure 8: Results of the high-resolution wind and temperature field reconstruction for case 100(3) u_{10} = 4 m/s, H_g = 0.2 m and S = 0%, depicting the high-resolution ground truth/target, mean prediction of the model, the reconstruction error, and comparison of the reconstructed value at the centre line of the field with model estimation and the model input.

Table 5
effect of increasing the number of input randomly sampled low-resolution versions on SR model accuracy utilising random sampling strategy and n=100

N. input samples	U* (m/s)		V* (m/s)		W* (m/s)		T* C	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
2	0.750	1.310	0.726	1.248	0.880	1.558	4.131	15.669
3	0.750	1.306	0.722	1.243	0.877	1.549	4.082	15.614
4	0.732	1.274	0.717	1.236	0.868	1.539	3.995	15.557

* U(m/s) ∈ [-18.7, 55.3], V (m/s)∈ [-38.4, 37.7], W (m/s)∈ [-22.2, 52.1]

high-resolution pairs extracted using all three strategies and varying UAV numbers. Consequently, by utilising three samples per UAV number for the second and third strategies, the training dataset includes 18 × 17325 input-output pairs of data, leaving the case with the third sampling strategy and n=49 out for testing the model. Besides, since the first strategy assumes UAVs are hovering at fixed locations, only a single sample is extracted for the scenario per UAV number. Finally, it is important to note that this model was trained for only 100 epochs due to the substantial computational time required and the observed lack of improvement after 40 epochs.

The model accuracy is evaluated utilising the average MAE, and RMSE of the model while the uncertainty of estimations is quantified utilising Std of accuracy metrics.

Table 6
performance evaluation metrics and uncertainty of the primary SR model against different testing subsets

N. UAV	U (m/s)		V (m/s)		W (m/s)		T	
(strategy)	MAE (std)	RMSE (std)	MAE (std)	RMSE (std)	MAE (std)	RMSE (std)	MAE (std)	RMSE (std)
25 (1)	0.755 (0.0095)	1.315 (0.0161)	0.731 (0.0088)	1.261 (0.0152)	0.883 (0.0106)	1.562 (0.0188)	4.158 (0.0501)	15.984 (0.1897)
49 (1)	0.752 (0.0092)	1.308 (0.0159)	0.730 (0.0090)	1.259 (0.0154)	0.881 (0.0106)	1.555 (0.0188)	4.111 (0.0494)	15.787 (0.1890)
100 (1)	0.748 (0.0094)	1.301 (0.0163)	0.729 (0.0088)	1.257 (0.0151)	0.878 (0.0107)	1.546 (0.0189)	4.052 (0.0496)	15.570 (0.1896)
25 (2)	0.762 (0.0087)	1.325 (0.0155)	0.732 (0.0088)	1.263 (0.0151)	0.884 (0.0106)	1.565 (0.0189)	4.189 (0.0495)	16.060 (0.1944)
49 (2)	0.756 (0.0097)	1.314 (0.0166)	0.731 (0.0088)	1.261 (0.0152)	0.882 (0.0106)	1.559 (0.0188)	4.145 (0.0499)	15.915 (0.1878)
100 (2)	0.752 (0.0096)	1.307 (0.0163)	0.730 (0.0088)	1.259 (0.0152)	0.880 (0.0106)	1.552 (0.0164)	4.097 (0.0434)	15.729 (0.1652)
25 (3)	0.759 (0.0075)	1.321 (0.0133)	0.732 (0.0077)	1.264 (0.0134)	0.884 (0.0093)	1.566 (0.0167)	4.189 (0.0434)	16.122 (0.1652)
100 (3)	0.755 (0.0088)	1.315 (0.0144)	0.732 (0.0077)	1.263 (0.0134)	0.883 (0.0093)	1.563 (0.0165)	4.164 (0.0440)	16.045 (0.1703)
Overall	**0.755 (0.0101)**	**1.313 (0.0177)**	**0.731 (0.088)**	**1.261 (0.0153)**	**0.882 (0.0108)**	**1.558 (0.0198)**	**4.138 (0.0736)**	**15.901 (0.2618)**

* U(m/s) ∈ [-18.7, 55.3], V (m/s)∈ [-38.4, 37.7], W (m/s)∈ [-22.2, 52.1]

For each testing subset, the model undergoes 15 inference cycles with randomly sampled low-resolution inputs to accompany each sample. Table 6 presents the average performance evaluation metrics of the SR model, along with the standard deviation of these metrics, evaluated against the testing sets of the low-resolution datasets used to train the model. It should be highlighted that since this section solely focuses on the model uncertainty, all inference cycles are performed on a fixed testing set for each scenario so that the data uncertainty does not affect the evaluations.

Overall, the model achieved an MAE and RMSE equal to 0.755, 1.313 (m/s) for the U component of velocity which ranges from -18.7 to 55.3 m/s. For the V component, which ranges from -38 to 37 m/s the MAE and RMSE were 0.731 and 1.261. The MAE and RMSE for the W component that ranges from -22.2 to 52.1 are equal to 0.882, and 1.558, while for temperature (T), they were 4.138 and 15.901. Generally, it could be concluded that the model presents acceptable accuracy based on a comparison of MAE and RMSE with the value ranges. Furthermore, the standard deviation of all metrics are bound within a range of 1.5% of the metrics, indicating a low uncertainty of model estimations within the training range of data.

In terms of the impact of the sampling strategy on the model output accuracy, while the first sampling strategy demonstrated a slightly lower MAE and RMSE, the differences between the three strategies were minimal and hence negligible, suggesting a proper model generalisation. In addition, increasing the number of UAVs from 25 to 100 also proved to have a negligible effect on the model output accuracy. In terms of the estimation uncertainty, however, it could be concluded that although indicating a marginally higher MAE and RMSE, the model exhibits a lower epistemic uncertainty in its estimation while the low-resolution input is obtained through the third sampling strategy. This outcome is however advantageous as the third strategy aims to replicate and is the closest scenario to the actual distribution of the UAVs within the field. Figure 8 presents a graphical representation of the SR model target, mean predictions, and error values while reconstructing the high-resolution wind and temperature field for the with case $u_{10} = 4m/s$, H_g=0.2 m, and horizontal terrain (S=0 %). The low-resolution samples were generated from the scenario with 100 UAVs utilising the third sampling strategy (100(3)). The figure is organised into columns presenting the ground truth (first column), reconstructed field (second column), and the error values (third column), for the stream-wise component of velocity (U) 8a, transverse component of velocity (V) 8b, normal component of velocity (W) 8c, and the temperature field (T)8d. It could be observed that while the model prediction is a smoothed version of the true field, the overall performance of the model is satisfactory, capturing the general pattern of the wind field within the fire plume and surrounding area. The last column compares the ground truth, input, and the model estimated values at the centre line of the field. The model demonstrates accurate estimation of the upstream wind field as well as significantly improved estimations compared to the interpolated input values. Additionally, the narrow uncertainty bounds around estimations indicate the model's low uncertainty.

Table 7
Evaluation of performance metrics and uncertainty of the primary SR model against unseen UAV numbers

		N. UAVs (sampling strategy)				
		49 (3)	16 (3)	9 (3)	9 (2)	9 (1)
U	**MAE**	**0.755**	**0.762**	**0.777**	**0.786**	**0.767**
	RMSE	1.314	1.327	1.347	1.359	1.333
	PICP	0.107	0.123	0.142	0.144	0.134
	PINAW (%)	**0.485**	**0.568**	**0.688**	**0.719**	**0.661**
V	**MAE**	**0.732**	**0.733**	**0.734**	**0.734**	**0.733**
	RMSE	1.262	1.265	1.266	1.266	1.265
	PICP	0.057	0.059	0.061	0.063	0.060
	PINAW(%)	**0.168**	**0.176**	**0.184**	**0.192**	**0.182**
W	**MAE**	**0.884**	**0.885**	**0.887**	**0.887**	**0.886**
	RMSE	1.564	1.569	1.572	1.571	1.570
	PICP	0.058	0.060	0.063	0.065	0.064
	PINAW(%)	**0.250**	**0.257**	**0.269**	**0.283**	**0.265**
T	**MAE**	**4.168**	**4.207**	**4.247**	**4.261**	**4.222**
	RMSE	16.052	16.185	16.274	16.211	16.198
	PICP	0.188	0.194	0.202	0.200	0.191
	PINAW(%)	**0.185**	**0.194**	**0.211**	**0.230**	**0.212**

4.1.3. *Effect of measurement uncertainties*

Data measurement uncertainty typically stems from two sources: measurement devices and the measurement points' spatial distribution, which can result in infinitely different sampling combinations —or low-resolution versions— of the same high-resolution field. Wind measurements using UAV swarms are typically subject to different error sources,

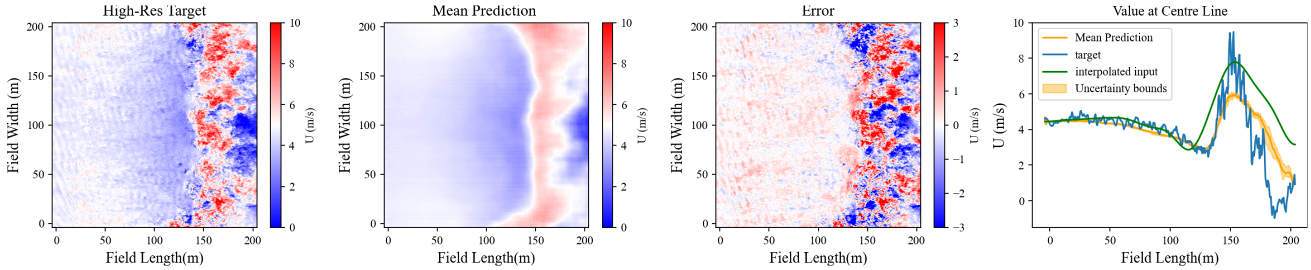
Table 8
Evaluation of performance metrics and uncertainty of the primary SR model against different measurement noise

		N. UAVs (sampling strategy)								
		9(2)				25(3)	49(3)			
	noise level	no noise	10%	30%	50%	50%	no noise	10%	30%	50%
U	**MAE**	**0.786**	**0.936**	**1.513**	**2.038**	**1.736**	**0.755**	**0.790**	**1.143**	**1.579**
	RMSE	1.359	1.530	2.138	2.730	2.369	1.314	1.358	1.743	2.190
	PICP	0.144	0.259	0.494	0.512	0.451	0.107	0.168	0.394	0.430
	PINAW(%)	**0.719**	**1.462**	**3.765**	**5.121**	**4.006**	**0.485**	**0.865**	**2.533**	**3.516**
V	**MAE**	**0.734**	**0.736**	**0.750**	**0.766**	**0.757**	**0.732**	**0.733**	**0.738**	**0.750**
	RMSE	1.266	1.270	1.293	1.321	1.304	1.262	1.264	1.273	1.280
	PICP	0.063	0.073	0.110	0.132	0.112	0.057	0.062	0.088	0.103
	PINAW(%)	**0.192**	**0.228**	**0.363**	**0.456**	**0.378**	**0.168**	**0.187**	**0.271**	**0.336**
W	**MAE**	**0.887**	**0.890**	**0.911**	**0.935**	**0.921**	**0.884**	**0.884**	**0.894**	**0.912**
	RMSE	1.571	1.579	1.628	1.676	1.650	1.564	1.566	1.588	1.628
	PICP	0.065	0.075	0.111	0.134	0.115	0.058	0.064	0.091	0.107
	PINAW(%)	**0.283**	**0.335**	**0.524**	**0.656**	**0.553**	**0.251**	**0.278**	**0.404**	**0.501**
T	**MAE**	**4.261**	**4.501**	**5.263**	**5.937**	**5.523**	**4.168**	**4.233**	**4.714**	**5.292**
	RMSE	16.211	16.381	17.016	17.624	17.274	16.052	16.094	16.418	16.913
	PICP	0.200	0.266	0.431	0.440	0.376	0.188	0.207	0.356	0.351
	PINAW(%)	**0.230**	**0.327**	**0.621**	**0.818**	**0.666**	**0.185**	**0.242**	**0.444**	**0.595**

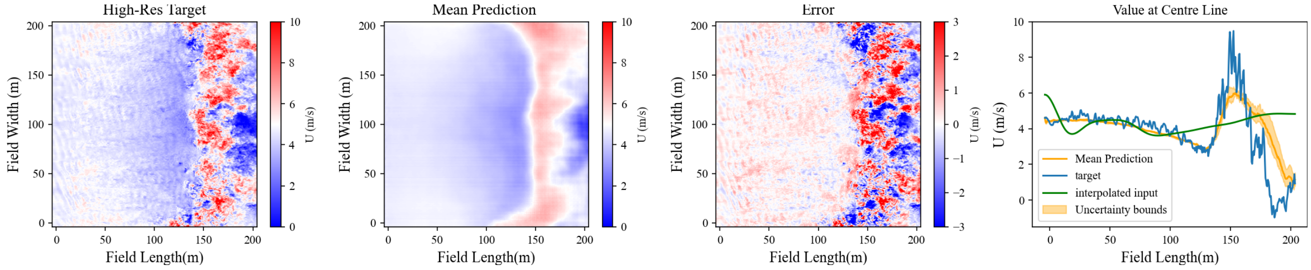
leading to high levels of uncertainty. Additionally, the number of UAVs in real-world scenarios may be significantly lower than those considered in our simulations. Therefore, the robustness and performance of the trained model are further investigated against the unseen number of UAVs, as well as various levels of noise in the measured wind values. Contrary to the previous section where the model was inferenced multiple times with a fixed dataset, in this section, each high-resolution wind field within the test subset, is randomly sampled multiple times to generate several Low-resolution versions of the same high-resolution wind field, replicating the infinitely possible spatial distribution of UAVs over the domain. These versions are then used to make multiple estimations of the same high-resolution field, capturing the combined effect of model uncertainty and data uncertainty. As a result, the inference time and complexity of experiments in this section are significantly higher than those of the previous section. Table 7 presents the accuracy and uncertainty metrics of the model, while evaluated against previously unseen combinations of the number of UAVs and sampling strategies.

The first experiment quantifies the effect of decreasing the number of UAVs from 49 to 16 and subsequently to 9 while utilising third sampling strategies. Generally, it is indicated that even though the model is not trained on data from the case with 49 UAVs and third sampling strategy 49(3), it provides equally accurate estimations as those seen during the training phase, i.e. achieving an average MAE =0.755 and an RMSE=1.314 for U component of velocity. Further reducing the number of UAVs from 49 to 9 leads to a slight increase of MAE from 0.755 to 0.777, which corresponds to 3% increase. However, this increase is relatively minor, suggesting that the model can provide reliable estimations despite the low-resolution input field being significantly altered by the reduced number of sampling points. The model uncertainty is however quantified in terms of Prediction Interval Coverage Probability (PICP), and the Prediction Interval Normalized Average Width (PINAW) parameters. This is mainly due to the minimal variations of the MAE and RMSE standard deviations across different cases, making it difficult to identify trends based on that alone. As previously mentioned, a higher PICP indicates more target values are covered by the predicted prediction intervals (PIs), however, this does not necessarily imply a better model. Considering the accuracy and uncertainty metrics presented in Table 7, it can be observed that the PICP for the U component of velocity increases by decreasing the number of UAVs from case 49(3) to 9(3). Simultaneously, the averaged width of prediction intervals (PINAW) increased from 0.485% to 0.688%, equivalent to 42% increase, as the number of UAVs decreased from 49 to 9, suggesting that the higher PICP in cases with smaller UAV numbers primarily corresponds to the increased width of PIs, reflecting higher uncertainty of the model predictions along with the slightly reduced average accuracy of the model. This is however more pronounced in the U and T parameters, where the variations in PICP and PINAW are more significant. In contrast, it is less pronounced in the V and W parameters, as indicated by the smaller variations in their corresponding PICP and PINAW values. Comparing the PICP and PINAW metrics in cases with 9 UAVs and different sampling strategies, it could be observed that the average accuracy and uncertainty measures of the model for all cases are relatively similar with negligible variations of up to 3% for MAE and 0.002 and 0.01 in PICP and PINAW.

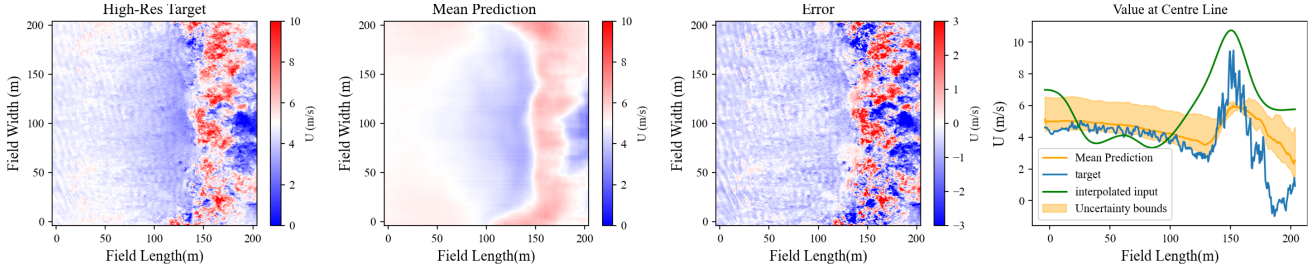
The second experiment revolves around quantifying the effect of measurement noise and uncertainty on the model's output accuracy and reliability. The model performance is hence evaluated using low-resolution inputs subjected to varying levels of measurement noise, replicating the uncertainties and errors that are likely to happen in real-world scenarios and environments. Similar to the previous experiment, multiple randomly sampled low-resolution versions of every high-resolution instant are generated. In this case, sampled data are subjected to random Gaussian noise and are then used to generate the low-resolution input file. This



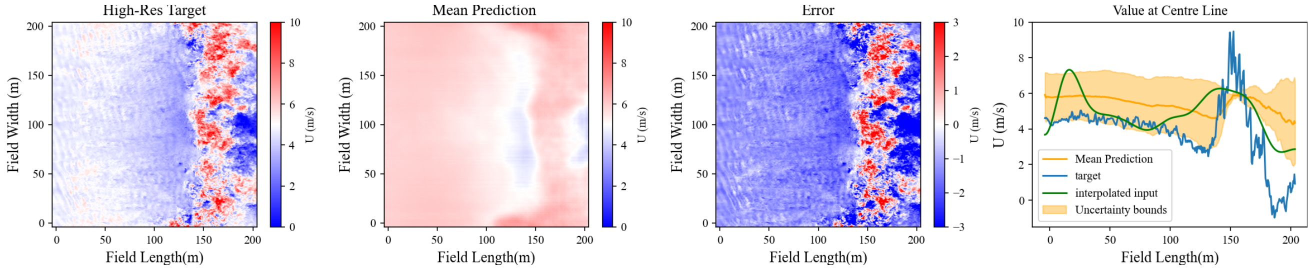
(a) Measurement noise= 0%



(b) Measurement noise= 10%



(c) Measurement noise= 30%



(d) Measurement noise= 50%

Figure 9: Results of the high-altitude reconstruction of the U-component for case 49(3), $u_{10} = 4$ m/s, $H_g = 0.2$ m and $S = 0\%$ under different measurement noise levels, depicting the high-resolution ground truth/target, mean prediction of the model, the reconstruction error, and comparison of the reconstructed value at the centre line of the field with model estimation and the model input.

approach mimics the infinitely possible spatial distributions of UAVs, as well as the potential errors of mounted sensors. The input Gaussian noise (N) is defined as a function of the input data standard deviation for each velocity component and each input batch which results in the noisy input data in the form of $U_{Noisy} = N \times U + U$. The obtained accuracy metrics for the test dataset with the prescribed noise levels including 10%, 30% and 50% are presented in Table 8.

Three different combinations of UAV number and sampling strategy are considered. The obtained results indicate that generally, increasing the measurement noise level decreases the model's output accuracy, and increases its uncertainty. For instance, in the case with 9 UAVs using third sampling strategy 9(3), introducing only 10% measurement noise has resulted in approximately 19% increase in MAE, and 12% in MSE corresponding to the U component of velocity, and 6% and 1% for the corresponding metrics concerning the the temperature. However, for V and W, changes are mainly limited to 1%. In addition, the model uncertainty is also increased as indicated through increased normalised average width of PI (PINAW) values, from 0.719% to 1.462% for the U, from 0.192 to 0.228% for the V, from 0.283 to 0.335% for W, and from 0.230 to 0.327% for temperature (T).

Table 9
performance evaluation metrics and uncertainty of the near surface model against different testing subsets

	U (m/s)		V (m/s)		W (m/s)	
UAV number (strategy)	MAE (std)	RMSE (std)	MAE (std)	RMSE (std)	MAE (std)	RMSE (std)
perfect data	0.744	1.202	0.627	1.071	0.545	1.004
25 (1)	0.862	1.337	0.679	1.142	0.552	1.011
49 (1)	0.859	1.335	0.677	1.14	0.551	1.01
100 (1)	0.857	1.331	0.676	1.138	0.551	1.01
25 (2)	0.867	1.345	0.679	1.142	0.552	1.01
49 (2)	0.862	1.337	0.678	1.141	0.552	1.01
100 (2)	0.859	1.334	0.677	1.139	0.551	1.01
25 (3)	0.867	1.344	0.68	1.144	0.552	1.011
100 (3)	0.864	1.341	0.679	1.142	0.551	1.01
Overall	0.849 (0.040)	1.323 (0.046)	0.672 (0.017)	1.022 (0.023)	0.551 (0.002)	1.010 (0.002)

* U(m/s) ∈ [-26.1, 41.3], V (m/s)∈ [-33.1, 27.35], W (m/s)∈ [-25.6, 29.95]

Concerning the impact of varying numbers of UAVs, it could be concluded that typically increasing the number of UAVs improves the model accuracy and reduces the model uncertainty under all input noise conditions. However, this effect is more pronounced when input noise levels are higher. For example, increasing the number of UAVs from 9 to 25 under 50% measurement noise conditions leads to a reduction of the MAE and PINAW values from 2.038 to 1.736 and from 5.121 to 4.006, representing a 15% reduction in MAE and 22% reduction in PINAW values, respectively.

Figure 9 presents a graphical representation of the super-resolution reconstruction of the U component of velocity with 49 UAVs and utilising the third sampling strategy, under different measurement noise levels. The environmental conditions of this instant include $u_{10} = 4m/s$, H_g=0.2 m, and horizontal terrain (S=0 %). Generally, it could be observed that the quality and accuracy of reconstruction reduce as the input noise increases. Increased width of PIs corresponding to the increased levels of input noise are also evidently presented. Notably, by increasing the input noise level, the average estimations of the velocity tend to overestimate the wind velocity at regions behind the fire plume, resulting in a smoother and more uniform estimation of the wind field estimation under high levels of noise. This, however, holds while $u_{10} \leqslant 6$ m/s. Nevertheless, for higher wind velocities, increasing the input noise leads to an underestimation of the wind velocity, suggesting that the model outputs begin to converge to a mean flow field. For the temperature field (Figure S1.1), the model consistently underpredicts temperature values within the plume across varying noise levels. However, it indicates a similar trend to the velocity components for temperature values ahead of the fire plume, with the difference in predictions being small and mainly limited to 0.5 degrees.

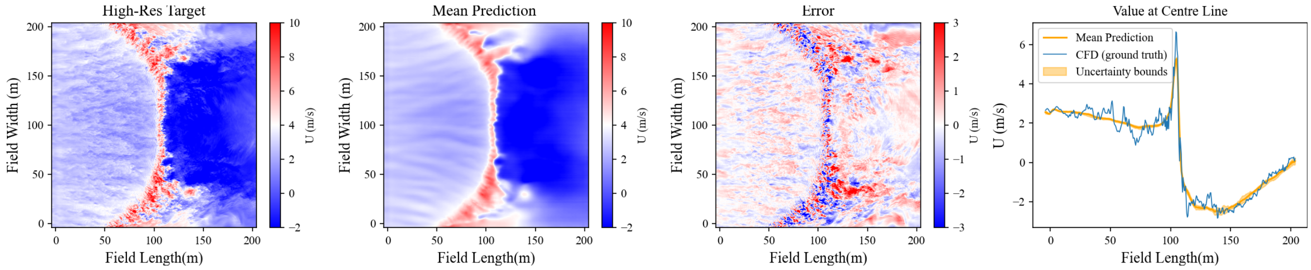
4.2. Near surface wind field estimation

This section presents the accuracy metrics and uncertainty measures obtained from models trained to estimate the near-surface wind field. As previously described, the purpose of this section is to train a model such that it estimates the near-surface wind field velocity components $I_{\text{surf}} = (u_2, v_2, w_2) \in \mathbb{R}^{416\times416\times3}$ from a concatenation of reconstructed high altitude wind field -from previous section- and the high-resolution data corresponding to the surface temperature, vegetation height and type, and terrain slope $C = \text{concat}(I_{sr}, I_{\text{terrain}}) \in \mathbb{R}^{416\times416\times8}$. Similar to the previous section, the effect of embedding the attention mechanism in the model on the estimation accuracy is studied.

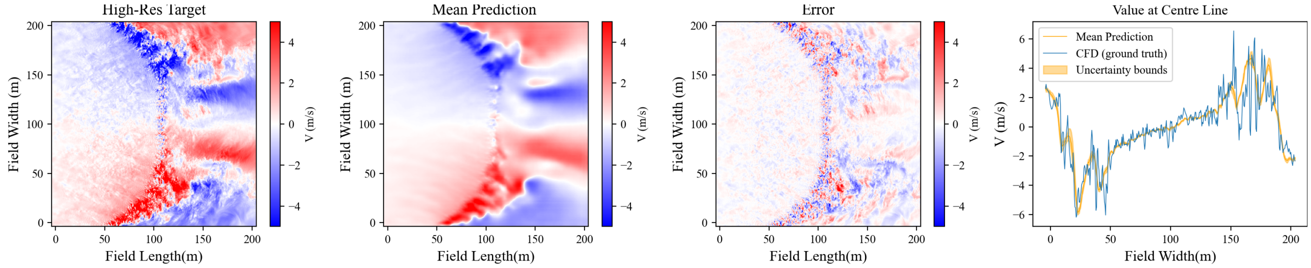
4.2.1. *Model performance and epistemic uncertainty*

The obtained results indicate that the fully convolutional model was able to provide accurate estimations of the wind velocity with MAE=0.785 and RMSE =1.255 for U, MAE=0.656 and RMSE =1.123 for V, and MAE=0.555 and RMSE =1.021 for W component of velocity, which are however up to 5.5% higher than corresponding values of the attention-based model presented in Table 9. It is also worth mentioning that the results generated by the fully convolutional model provide a smoother representation of the wind field compared to when the attention mechanism is utilized. Consequently, the attention mechanism is embedded into the final model. Table 9 presents the average performance evaluation metrics for the near surface wind estimation model evaluated against the testing subsets of the reconstructed wind fields used to train the model. Similar to the section 4.1.2, the model has undergone multiple inference cycles for each testing subset with randomly sampled batches to accompany each sample. Overall, the model achieved MAE and RMSE equal to 0.849 and 1.323 for the U, 0.672 and 1.142 for the V, and 0.551 and 1.01 for the W velocity component, respectively. Furthermore, the standard deviation of all metrics is bound in a range of 5% of the metrics, indicating relatively high confidence in model estimations and low epistemic uncertainty.

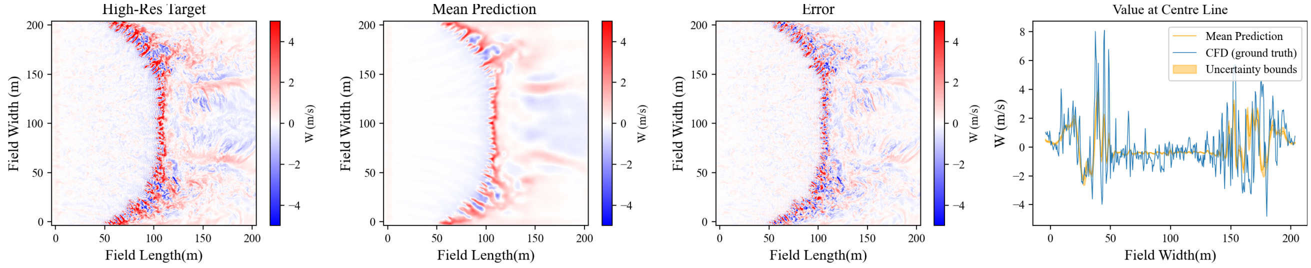
Even though the third sampling strategy shows a slightly higher MAE and RMSE compared to the other strategies, the impact of the high-altitude wind sampling strategy proved to be negligible as indicated by the minor variation in MAE and RMSE metrics across different scenarios. Changing the number of UAVs from 25 to 100 on the near-surface wind field estimation accuracy described a negligible effect of less than 1% on the average accuracy of the model across different scenarios as presented in Table 9, thus it could be considered negligible. Figure 10 shows the estimated velocity components (second column), ground truth (first



(a) Estimation results for U (m/s) component of velocity



(b) Estimation results for V (m/s) component of velocity



(c) Estimation results for W (m/s) component of velocity

Figure 10: 49 (3) near surface-no noise-sample 16

column), error values ($y_i - y_i'$, third column), and a comparison of the estimated and ground truth values along the centerline of the field (fourth column) for the case of $u_{10} = 4, \text{m/s}$, $H_g = 0.2, \text{m}$, and flat terrain ($S = 0\%$) at a specific time instance. The near-surface wind is estimated utilizing the reconstructed high-altitude wind field from the scenario with 49 UAVs utilising the third sampling strategy (49(3)). Overall, the model demonstrates a satisfactory performance while estimating the near-surface wind fields based the corresponding reconstructed instantaneous wind field at 40m *AGL* as depicted in Figure 10a, 10b, and 10c for U, V, and W component of velocity. Testing the model across various wind conditions and terrain slopes shows that it has successfully learned an accurate mapping between input and output values. The visualization of the errors indicates that the majority of the inaccuracies stem from the stochastic and instantaneous nature of the wind field, which complicates the estimation of its oscillatory behaviour. Overall, it can be concluded that the general trend is captured and estimated with satisfactory accuracy.

The results for the estimation of the U component of velocity -as depicted in Figures 10a, S1.4a, and S1.7a- indicate that the model accurately captures the span and formation of the fire-induced wind field, including the accelerated wind above the flaming front and the downstream vortex zone. The estimations of the V-velocity component -as depicted in 10b, S1.5a, and S1.8a- show that the model estimates the maximum and minimum values fairly well, particularly within the downwind turbulent structure of the fire-induced wind. However, the W component of velocity is highly oscillatory -as depicted in 10c, S1.6a, and S1.9a-, leading to higher estimation error values in certain areas, specifically within the downstream vortex zone. It is important to note

Table 10
Evaluation of performance metrics and uncertainty of the near surface model against unseen UAV numbers

		N. UAVs (sampling strategy)				
		49 (3)	16 (3)	9 (3)	9 (2)	9 (1)
U (m/s)	**MAE**	**0.861**	**0.873**	**0.880**	**0.884**	**0.873**
	RMSE	1.337	1.352	1.362	1.367	1.353
	PICP	0.0597	0.0653	0.0708	0.0725	0.0689
	PINAW (%)	**0.6410**	**0.7210**	**0.7960**	**0.8270**	**0.783**
V (m/s)	**MAE**	**0.679**	**0.681**	**0.685**	**0.683**	**0.682**
	RMSE	1.142	1.146	1.151	1.149	1.146
	PICP	0.075	0.078	0.080	0.081	0.07920
	PINAW (%)	**0.449**	**0.472**	**0.489**	**0.502**	**0.48300**
W(m/s)	**MAE**	**0.552**	**0.552**	**0.553**	**0.552**	**0.552**
	RMSE	1.011	1.011	1.012	1.011	1.011
	PICP	0.047	0.048	0.049	0.049	0.049
	PINAW (%)	**0.2710**	**0.2770**	**0.2830**	**0.2860**	**0.2830**

Table 11
Evaluation of performance metrics and uncertainty of the near surface model against different measurement noise

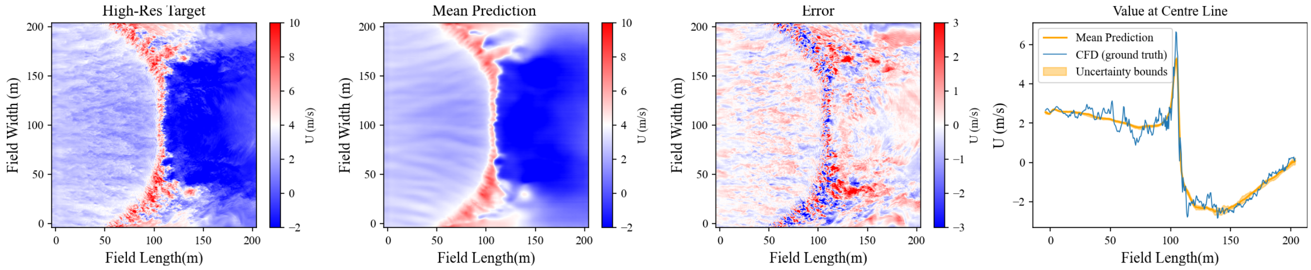
		N. UAVs (sampling strategy)									
		9(2)				25(3)		49(3)			
	Noise level	0 %	10%	30%	50%	0 %	50%	0 %	10%	30%	50%
U (m/s)	**MAE**	**0.884**	**0.965**	**1.123**	**1.490**	**0.867**	**1.358**	**0.861**	**0.893**	**1.057**	**1.267**
	RMSE	1.367	1.460	1.742	2.073	1.344	1.875	1.337	1.374	1.544	1.766
	PICP	0.073	0.107	0.187	0.205	0.063	0.179	0.060	0.074	0.148	0.166
	PINAW (%)	**0.827**	**1.283**	**2.642**	**3.479**	**0.687**	**2.734**	**0.641**	**0.837**	**1.798**	**2.384**
V (m/s)	**MAE**	**0.683**	**0.690**	**0.710**	**0.761**	**0.68**	**0.748**	**0.679**	**0.681**	**0.702**	**0.735**
	RMSE	1.149	1.157	1.197	1.255	1.144	1.235	1.142	1.145	1.171	1.213
	PICP	0.081	0.096	0.137	0.156	0.0768	0.139	0.075	0.081	0.114	0.130
	PINAW (%)	**0.5020**	**0.585**	**0.862**	**1.026**	**0.461**	**0.901**	**0.4490**	**0.488**	**0.693**	**0.831**
W(m/s)	**MAE**	**0.552**	**0.554**	**0.558**	**0.565**	**0.552**	**0.563**	**0.552**	**0.552**	**0.556**	**0.561**
	RMSE	1.011	1.012	1.016	1.022	1.011	1.019	1.011	1.011	1.014	1.017
	PICP	0.049	0.057	0.078	0.088	0.0475	0.078	0.047	0.050	0.066	0.073
	PINAW (%)	**0.286**	**0.316**	**0.401**	**0.451**	0.275	**0.414**	**0.271**	**0.286**	**0.351**	**0.392**

that the W component is inherently the most oscillatory, making it the most challenging to estimate. However, it should be noted that in certain areas—particularly near the flaming front, where maximum velocity values occur—the estimates are less accurate and can be underestimated up to 24% for the U component of velocity and up to 60% for the W component. This is however not surprising as the model is trained to estimate the average and large-scale behaviour of the flow rather than the oscillatory and vortical structure of the wind field at small scales from sparse measurements carried out at the UAVs' flying altitude.

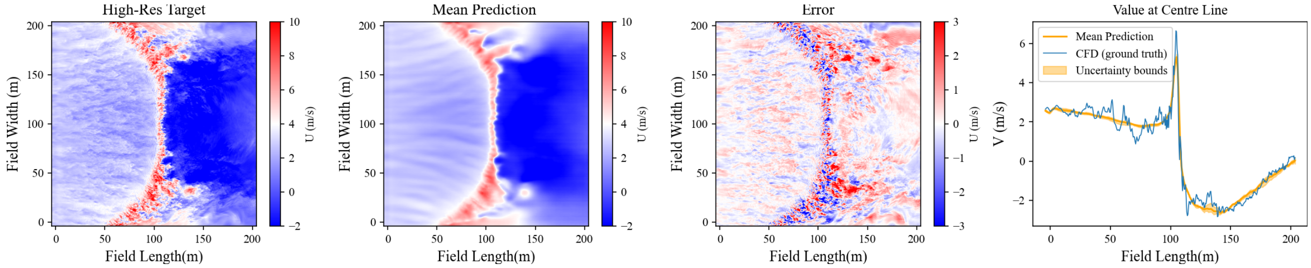
4.2.2. Effect of measurement uncertainty

This section examines the effect of data measurement uncertainties including the spatial distribution of sampling points and the measurement device errors, on model performance and estimation reliability. As discussed in Section 4.1.3, multiple low-resolution versions were generated for each high-resolution wind field at flying altitude, which were then used to reconstruct the high-altitude wind field and conduct uncertainty analysis. These reconstructed high-altitude wind fields are subsequently utilized in this section to generate multiple representations of the near-surface wind field, providing both average estimations and uncertainty measures. Consequently, the uncertainty propagates throughout the entire process, as the uncertain estimations of the first model are used to generate uncertain near-surface wind field estimations in the second model, thereby accounting for the total uncertainty of the approach. Table 10 presents the accuracy metrics and uncertainty measures of the model, against unseen combinations of the number of UAVs and sampling strategies. The results indicate that the model is generally robust to changes in the number of sampling points or UAVs. Reducing the number of UAVs from 49 to 9 led to a slight increase in the MAE, corresponding to an approximate rise of 2.5% from 0.86 to 0.88 for the U component of velocity. The RMSE metric also shows a similar trend, with a slight increase, while the uncertainty metric, PINAW, grows from 0.6410 to 0.7960 as UAV numbers decrease, indicating a slightly wider range of uncertainty. However, the model indicates consistency in terms of accuracy metrics and uncertainty measures for the V and W components of velocity, regardless of the number of UAVs. As a result, it can be concluded that the model maintains high accuracy across various UAV configurations with a slight increase in uncertainty as the number of UAVs decreases.

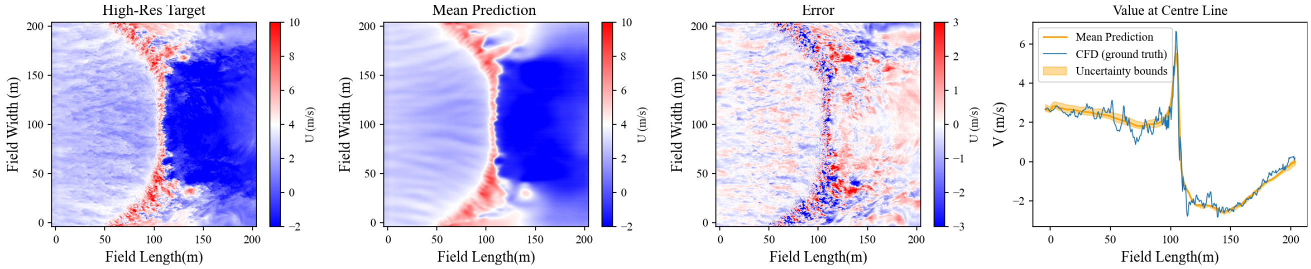
Table 11 presents the evaluation of the near-surface estimations of the second model under different levels of measurement noise. The uncertainty caused by measurement noise propagates throughout the entire process, as multiple noisy low-resolution measurements are utilised to generate several versions of high-altitude wind fields. These in turn are subsequently used to generate multiple representations of the near-surface wind field, providing both average estimations and uncertainty measures. As expected, increased measurement noise leads to reduced estimation accuracy and higher uncertainty especially for smaller numbers of UAVs. For example, increasing the input noise to 50% leads to a rise in MAE from 0.884 to 1.49 for case 9(2), corresponding to 69% increase, which demonstrates a clear sensitivity to input noise. As noise increases, RMSE values also rise significantly, indicating larger discrepancies between predicted and true values. Both PICP and PINAW measures increase by increasing the input noise, particularly for case 9(2) where PICP for U increases from 0.073 to 0.205 (180%) and PINAW rises from 0.827 to 3.479 (320%). Considering the effect of UAV numbers, it could be concluded that increasing the number of UAVs, enhances the model robustness against input noise, as evident from the lower MAE, RMSE, PICP, and PINAW values for case 49(3) compared to 25(3) and 9(2). Regarding the V and W components of velocity, it could be highlighted that while uncertainty generally increases with noise -for example, PINAW values change from 0.502 to 1.026 for V and from 0.286 to 0.316 for W in case 9(2)-, this increase is less dramatic compared to the U component. This is however partly due to the fact that, despite their broad range, a significant portion of the V and especially W component values are near or equal to zero. When the model predicts these small values, the uncertainty has little impact on the estimations, resulting in lower average uncertainty metrics. Even as the uncertainty increases in regions with non-zero values, the overall average remains



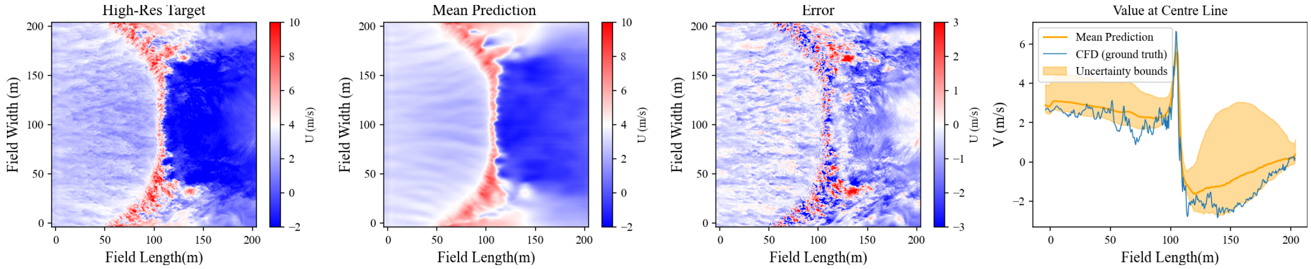
(a) Measurement noise= 0%



(b) Measurement noise= 10%



(c) Measurement noise= 30%



(d) Measurement noise= 50%

Figure 11: Results of the near-surface U-component of wind velocity estimation for case 49(3), $u_{10} = 4$ m/s, $H_g = 0.2$ m and $S = 0\%$, under different measurement noise levels, showing the high-resolution ground truth/target, the model's mean prediction, the reconstruction error, and a comparison of the estimated value along the centerline of the field with the ground truth (CFD)

small because many of the predicted values are small or close to zero.

Figure 11 presents the estimated U component of the near-surface wind field for case 49(3), $u_{10} = 4$ m/s, $H_g = 0.2$ m and $S = 0\%$, under different measurement noise levels. Other examples of wind field estimations are also presented in Figure S1.4 and S1.7 for the U component, Figure S1.2, S1.5, and S1.8 for the V component, and Figure S1.3, S1.6, and S1.9 for the W component of velocity. It can generally be observed that although the model produces a satisfactorily smoothed version of the wind field, the quality and accuracy of the estimations decline as input noise increases. This is reflected in the widening of prediction intervals (PIs), which corresponds to the higher levels of input noise. The increased uncertainty in the predictions highlights the sensitivity of the model to noisy input data, leading to less precise estimations under higher noise conditions.

5. Model Generalization and Limitations

This section briefly addresses the practical capabilities and current limitations of the proposed method. As noted, the model was trained and tested on a square terrain with a constant slope, simulating a line fire ignited at the centre of the upstream boundary. But how applicable is this method to a large-scale wildfire? We believe that this method could be generalized by adopting a patching technique, where the desired area or firefront is divided into a series of square patches and the wind field is estimated for each patch independently. Within each patch, the fire could be considered a line fire, making the proposed framework feasible for use in larger wildfires.

However, the model in its current form has a series of limitations. First, the terrain is unlikely to remain constant and uniform within a $200 \times 200\ m^2$ area. Besides, the assumption of a symmetric line fire across all patches seems to be unrealistic as fire fronts can vary significantly in shape within each patch. Another important limitation is that the model is trained with input wind directed along the field length, meaning that patches must be aligned parallel to the ambient wind which would be extremely difficult to achieve in actual fires and complex terrains. Finally, the model assumes that all data is collected at a constant height, directly parallel to the terrain, which may be difficult to obtain in actual fire conditions.

These drawbacks point to opportunities where further study can improve the system and make it more flexible for use in real-world situations. The proposed technique has the potential to be useful for research projects as well as fire management and suppression tactics with additional development.

6. Conclusions

This study suggests a novel framework based on a combination of real-time UAV-swarm-based measurements of the fire-induced wind field and the utilization of Deep learning techniques for downscaling the measured wind field at the measurement height to the near-surface wind field at the desired spatial resolution. A series of 3D simulations of the grassland fire propagation is performed using the WFDS software in different configurations of wind speed and terrain slope in order to provide a dataset of synthetic data for training the DL model. The wind field at two different heights of 40m *AGL* and 2m *AGL* are then extracted from the model at different intervals and are used as input and target values of the DL model. The utilized DL model includes an attention-based residual UNet which is trained in a supervised procedure in order to learn a mapping from the input wind field (40m *AGL*) onto the target wind field (2m *AGL*) at different stages of the fire propagation. The trained model is then examined against different input conditions and against different ratios of noise in the input wind field. Results indicate that the proposed framework is capable of providing fast and accurate estimations of the three-dimensional velocity components of the near-surface wind at 2m *AGL* from the measured wind field at 40m *AGL*. The trained model reached an overall *MAE* and RMSE of 0.321 m/s and 0.644 m/s, respectively, for the longitudinal component of velocity (U[-7.42, 31.46]), 0.219 and 0.499 for the transverse component of velocity (V[-20.41,19.83]), and 0.147 and 0.430 for the normal component of velocity (W[-17,32,12.90]) which indicates that the errors are not significant compared to the range of the velocities variations. It is also indicated that the model is robust against input noise and shows satisfactory performance against unseen cases even against completely unseen cases. Considering that the wind measurement using *UAV*s could include significant levels of noise, the model is then tested against different ratios of input Gaussian noise of up to 50%. The obtained results indicated that even though the output results become so noisy, the model is still capable of estimating the average behaviour of the wind field even against high levels of input noise.

7. Abbreviations

CRediT authorship contribution statement

Mohammad Tavakol Sadrabadi: Conceptualization, Methodology, Software, Writing, Investigation, Visualization. **Mauro Sebastián Innocente:** Conceptualization, Methodology, Software, Writing, Supervision, Resources, Visualization.

References

[1] John T. Abatzoglou and A. Park Williams. Impact of anthropogenic climate change on wildfire across western us forests. *Proceedings of the National Academy of Sciences*, 113(42):11770–11775, 2016.

[2] Nerilie J. Abram, Benjamin J. Henley, Alex Sen Gupta, Tanya J. R. Lippmann, Hamish Clarke, Andrew J. Dowdy, Jason J. Sharples, Rachael H. Nolan, Tianran Zhang, Martin J. Wooster, Jennifer B. Wurtzel, Katrin J. Meissner, Andrew J. Pitman, Anna M. Ukkola, Brett P. Murphy, Nigel J. Tapper, and Matthias M. Boer. Connections of climate change and variability to large and extreme forest fires in southeast australia. *Communications Earth and Environment*, 2(1):8, 2021.

[3] Brian E. Potter. Atmospheric interactions with wildland fire behaviour – i. basic surface interactions, vertical profiles and synoptic structures. *International Journal of Wildland Fire*, 21(7):779–801, 2012.

[4] Brian E. Potter. Atmospheric interactions with wildland fire behaviour – 2. plume and vortex dynamics. *International Journal of Wildland Fire*, 21(7):802–817, 2012.

[5] S. B. Show. Climate and forest fires in northern california. *Journal of Forestry*, 17(8):965–979, 1919.

[6] H. T. Gisborne. Measuring forest-fire danger in northern idaho. Report, 1928.

[7] Naian Liu, Jiao Lei, Wei Gao, Haixiang Chen, and Xiaodong Xie. Combustion dynamics of large-scale wildfires. *Proceedings of the Combustion Institute*, 38(1):157–198, 2021.

[8] J.-L. Dupuy, R. R. Linn, V. Konovalov, F. Pimont, J. A. Vega, and E. Jiménez. Exploring three-dimensional coupled fire–atmosphere interactions downwind of wind-driven surface fires and their influence on backfires using the higrad-firetec model. *International Journal of Wildland Fire*, 20(6):734–750, 2011.

[9] Jiawei Zhang, Marwan Katurji, Peyman Zawar-Reza, and Tara Strand. The role of helicity and fire–atmosphere turbulent energy transport in potential wildfire behaviour. *International Journal of Wildland Fire*, 32(1):29–42, 2023.

[10] M. Tavakol Sadrabadi, Mauro S. Innocente, E.I. Gkanas, and I. Papagiannis. Comparison of the effect of one-way and two-way fire-wind coupling on the modelling of wildland fire propagation dynamics, 2022.

[11] B J Stocks, M E Alexander, and R A Lanoville. Overview of the international crown fire modelling experiment (icfme). *Canadian Journal of Forest Research*, 34(8):1543–1547, 2004.

[12] Marwan Katurji, Bob Noonan, Jiawei Zhang, Andres Valencia, Benjamin Schumacher, Jessica Kerr, Tara Strand, Grant Pearce, and Peyman Zawar-Reza. Atmospheric turbulent structures and fire sweeps during shrub fires and implications for flaming zone behaviour. *International Journal of Wildland Fire*, 32(1):43–55, 2023.

[13] Ajinkya Desai, Scott Goodrick, and Tirtha Banerjee. Investigating the turbulent dynamics of small-scale surface fires. *Scientific Reports*, 12(1):10503, 2022.

[14] Jasmine Innocent, Duncan Sutherland, Nazmul Khan, and Khalid Moinuddin. Physics-based simulations of grassfire propagation on sloped terrain at field scale: motivations, model reliability, rate of spread and fire intensity. *International Journal of Wildland Fire*, pages –, 2023.

[15] A. M. G. Lopes, L. M. Ribeiro, D. X. Viegas, and J. R. Raposo. Simulation of forest fire spread using a two-way coupling algorithm and its application to a real wildfire. *Journal of Wind Engineering and Industrial Aerodynamics*, 193:103967, 2019.

[16] Marcos Vanella, Kevin McGrattan, Randall McDermott, Glenn Forney, William Mell, Emanuele Gissi, and Paolo Fiorucci. A multi-fidelity framework for wildland fire behavior simulations over complex terrain. *Atmosphere*, 12(2):273, 2021.

[17] A. Bakhshaii and E.A. Johnson. A review of a new generation of wildfire–atmosphere modeling. *Canadian Journal of Forest Research*, 49(6):565–574, 2019.

[18] Rodman Linn, Jon Reisner, Jonah J. Colman, and Judith Winterkamp. Studying wildfire behavior using firetec. *International Journal of Wildland Fire*, 11(4):233–246, 2002.

[19] William E. Mell, Mary Ann Jenkins, Jim S. Gould, and P. B. Cheney. A physics-based approach to modelling grassland fires. *International Journal of Wildland Fire*, 16:1–22, 2007.

[20] Mark A. Finney. Farsite: Fire area simulator-model development and evaluation, 1998.

[21] Patricia L. Andrews. Current status and future needs of the behaveplus fire modeling system. *International Journal of Wildland Fire.*, 23:21–33, 2014.

[22] Paolo; Grasso and Mauro. Innocente. A two-dimensional reaction-advection-diffusion model of the spread of fire in wildlands. In *Advances in Forest Fire Research*, pages 334–342.

[23] Paolo Grasso and Mauro S. Innocente. Physics-based model of wildfire propagation towards faster-than-real-time simulations. *Computers & Mathematics with Applications*, 80(5):790–808, 2020.

[24] Thayjes Srivas, Tomàs Artés, Raymond A. de Callafon, and Ilkay Altintas. Wildfire spread prediction and assimilation for farsite using ensemble kalman filtering1. *Procedia Computer Science*, 80:897–908, 2016.

[25] Li Tan, Raymond A. de Callafon, Jessica Block, Daniel Crawl, Tolga Çağlar, and Ilkay Altıntaş. Estimation of wildfire wind conditions via perimeter and surface area optimization. *Journal of Computational Science*, 61:101633, 2022.

[26] N. S. Wagenbrenner, J. M. Forthofer, B. K. Lamb, K. S. Shannon, and B. W. Butler. Downscaling surface wind predictions from numerical weather prediction models in complex terrain with windninja. *Atmospheric Chemistry and Physics*, 16(8):5229–5241, 2016. ACP.

[27] Mauro S. Innocente and Paolo Grasso. Self-organising swarms of firefighting drones: Harnessing the power of collective intelligence in decentralised multi-robot systems. *Journal of Computational Science*, 34:80–101, 2019.

[28] Moulay A. Akhloufi, Andy Couturier, and Nicolás A. Castro. Unmanned aerial vehicles for wildland fires: Sensing, perception, cooperation and assistance. *drones*, 5(1):15, 2021.

[29] W. Thielicke, W. Hübert, U. Müller, M. Eggert, and P. Wilhelm. Towards accurate and practical drone-based wind measurements with an ultrasonic anemometer. *Atmospheric Measurement Techniques*, 14(2):1303–1318, 2021. AMT.

[30] N. Vasiljević, M. Harris, A. Tegtmeier Pedersen, G. Rolighed Thorsen, M. Pitter, J. Harris, K. Bajpai, and M. Courtney. Wind sensing with drone-mounted wind lidars: proof of concept. *Atmospheric Measurement Techniques*, 13(2):521–536, 2020. AMT.

[31] Kansuke Sasaki, Minoru Inoue, Tomoya Shimura, and Masato Iguchi. In situ, rotor-based drone measurement of wind vector and aerosol concentration in volcanic areas. *Atmosphere*, 12(3):376, 2021.

[32] T. Wetz, N. Wildmann, and F. Beyrich. Distributed wind measurements with multiple quadrotor unmanned aerial vehicles in the atmospheric boundary layer. *Atmospheric Measurement Techniques*, 14(5):3795–3814, 2021. AMT.

[33] S. Prudden, A. Fisher, M. Marino, A. Mohamed, S. Watkins, and G. Wild. Measuring wind with small unmanned aircraft systems. *Journal of Wind Engineering and Industrial Aerodynamics*, 176:197–210, 2018.

[34] Kevin Höhlein, Michael Kern, Timothy Hewson, and Rüdiger Westermann. A comparative study of convolutional neural network models for wind field downscaling. *Journal of Meteorological Applications*, 27(6):e1961, 2020.

[35] Jason M. Forthofer, Bret W. Butler, and Natalie S. Wagenbrenner. A comparison of three approaches for simulating fine-scale surface winds in support of wildland fire management. part i. model formulation and comparison against measurements. *International Journal of Wildland Fire*, 23(7):969–981, 2014.

[36] Jason M. Forthofer, Bret W. Butler, and Natalie S. Wagenbrenner. A comparison of three approaches for simulating fine-scale surface winds in support of wildland fire management. part i.an exploratory study of the effect of simulated winds on fire growth simulations. *International Journal of Wildland Fire*, 23(7):969–981, 2014.

[37] Jason Forthofer, Kyle Shannon, and Bret Butler. simulating diurnally driven slope winds with windninja. Report, 2009.

[38] James Hilton and Nikhil Garg. Rapid wind–terrain correction for wildfire simulations. *International Journal of Wildland Fire*, 30(6):410–427, 2021.

[39] D. A. Sachindra, F. Huang, A. Barton, and B. J. C. Perera. Statistical downscaling of general circulation model outputs to precipitation—part 2: bias-correction and future projections. *International Journal of Climatology*, 34(11):3282–3303, 2014.

[40] John T. Abatzoglou and Timothy J. Brown. A comparison of statistical downscaling methods suited for wildfire applications. *International Journal of Climatology*, 32(5):772–780, 2012.

[41] O. Rios, W. Jahn, E. Pastor, M. M. Valero, and E. Planas. Interpolation framework to speed up near-surface wind simulations for data-driven wildfire applications. *International Journal of Wildland Fire*, 27(4):257–270, 2018.

[42] Jincheng Zhang and Xiaowei Zhao. Spatiotemporal wind field prediction based on physics-informed deep learning and lidar measurements. *Applied Energy*, 288:116641, 2021.

[43] Jérôme Dujardin and Michael Lehning. Wind-topo: Downscaling near-surface wind fields to high-resolution topography in highly complex terrain with deep learning. 148(744):1368–1388, 2022.

[44] Karen Stengel, Andrew Glaws, Dylan Hettinger, and Ryan N. King. Adversarial super-resolution of climatological wind and solar data. *Proceedings of the National Academy of Sciences*, 117(29):16805, 2020.

[45] Jianxin Cheng, Jin Liu, Zhou Xu, Chenkai Shen, and Qiuming Kuang. Generating high-resolution climate prediction through generative adversarial network. *Procedia Computer Science*, 174:123–127, 2020.

[46] Kai Fukami, Koji Fukagata, and Kunihiko Taira. Super-resolution analysis via machine learning: a survey for fluid flows. *Theoretical and Computational Fluid Dynamics*, 37(4):421–444, August 2023.

[47] Kai Fukami, Koji Fukagata, and Kunihiko Taira. Super-resolution reconstruction of turbulent flows with machine learning. *Journal of Fluid Mechanics*, 870:106–120, 2019.

[48] Jinsong Zhang, Jian Liu, and Zhenwei Huang. Improved deep learning method for accurate flow field reconstruction from sparse data. *Ocean Engineering*, 280:114902, 2023.

[49] Pierre Dubois, Thomas Gomez, Laurent Planckaert, and Laurent Perret. Machine learning for fluid flow reconstruction from limited measurements. *Journal of Computational Physics*, 448:110733, 2022.

[50] Longyan Wang, Meng Chen, Zhaohui Luo, Bowen Zhang, Jian Xu, Zilu Wang, and Andy C.C. Tan. Dynamic wake field reconstruction

of wind turbine through physics-informed neural network and sparse lidar data. *Energy*, 291:130401, 2024.

[51] Moloud Abdar, Farhad Pourpanah, Sadiq Hussain, Dana Rezazadegan, Li Liu, Mohammad Ghavamzadeh, Paul Fieguth, Xiaochun Cao, Abbas Khosravi, U. Rajendra Acharya, Vladimir Makarenkov, and Saeid Nahavandi. A review of uncertainty quantification in deep learning: Techniques, applications and challenges. *Information Fusion*, 76:243–297, 2021.

[52] Adam M. Collins, Peter Rivera-Casillas, Sourav Dutta, Orie M. Cecil, Andrew C. Trautz, and Matthew W. Farthing. Super-resolution and uncertainty estimation from sparse sensors of dynamical physical systems. *Frontiers in Water*, 5, 2023.

[53] Yarin Gal and Zoubin Ghahramani. Dropout as a bayesian approximation: Representing model uncertainty in deep learning, 2016.

[54] Mattias Teye, Hossein Azizpour, and Kevin Smith. Bayesian uncertainty estimation for batch normalized deep networks, 2018.

[55] Mohammad Tavakol Sadrabadi and Mauro Sebastián Innocente. Enhancing wildfire propagation model predictions using aerial swarm-based real-time wind measurements: A conceptual framework. *Applied Mathematical Modelling*, 130:615–634, 2024.

[56] Geoffrey W. Donnell, Jordan A. Feight, Nate Lannan, and Jamey D. Jacob. Wind characterization using onboard IMU of sUAS. *2018 Atmospheric Flight Mechanics Conference*, 2018.

[57] Alexander Rautenberg, Martin S. Graf, Norman Wildmann, Andreas Platis, and Jens Bange. Reviewing wind measurement approaches for fixed-wing unmanned aircraft. *Atmosphere*, 9(11):422, 2018.

[58] Natapol Korprasertsak and Thananchai Leephakpreeda. Nyquist-based adaptive sampling rate for wind measurement under varying wind conditions. *Renewable Energy*, 119:290–298, 2018.

[59] NP Cheney, JS Gould, and WR Catchpole. The influence of fuel, weather and fire shape variables on fire-spread in grasslands. *International Journal of Wildland Fire*, 3(1):31–44, 1993.

[60] Charles R. Boardman, Mark A. Dietenberger, and David R. Weise. Specific heat capacity of wildland foliar fuels to 434 °c. *Fuel*, 292:120396, 2021.

[61] A. Kaiss B. Porterie, J. L. Consalvi and J. C. Loraud. Predicting wildland fire behavior and emissions using a fine-scale physical model. *Numerical Heat Transfer, Part A: Applications*, 47(6):571–591, 2005.

[62] Kevin McGrattan, Simo Hostikka, Jason Floyd, Randall McDermott, Marcos Vanella, and Eric Mueller. *Fire Dynamics Simulator User's Guide*, sixth edition, 2023.

[63] A.M. Grishin. *Mathematical Modeling of Forest Fires and New Methods of Fighting Them*. Publishing House of the Tomsk State University, 1997.

[64] Elham Amini, Mohammad-Saeed Safdari, Nathan Johnson, David R. Weise, and Thomas H. Fletcher. Pyrolysis kinetics of wildland vegetation using model-fitting methods. *Journal of Analytical and Applied Pyrolysis*, 157:105167, 2021.

[65] N. Boonmee and J.G. Quintiere. Glowing ignition of wood: the onset of surface combustion. *Proceedings of the Combustion Institute*, 30(2):2303–2310, 2005.

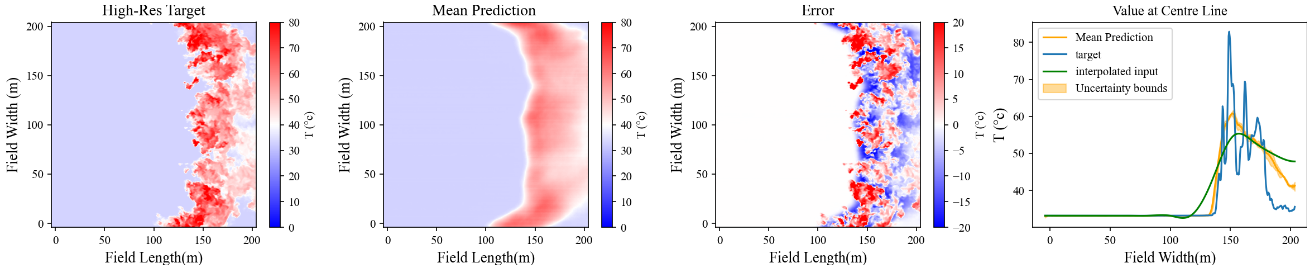
[66] N. Jarrin, J.-C. Uribe, R. Prosser, and D. Laurence. *Synthetic inflow boundary conditions for wall bounded flows*, volume 97 of *Notes on Numerical Fluid Mechanics and Multidisciplinary Design*. Springer Nature, United States, 2008.

[67] Jasmine Innocent, Duncan Sutherland, Nazmul Khan, and Khalid Moinuddin. Physics-based simulations of grassfire propagation on sloped terrain at field scale: motivations, model reliability, rate of spread and fire intensity. *International Journal of Wildland Fire*, 32:496–512, 2023.

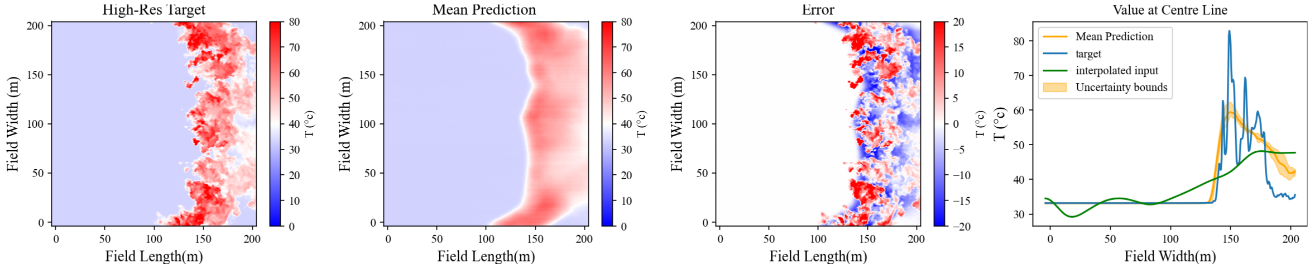
[68] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778.

[69] Sanghyun Woo, Jongchan Park, Joon-Young Lee, and In So Kweon. Cbam: Convolutional block attention module. Computer Vision – ECCV 2018, pages 3–19. Springer International Publishing.

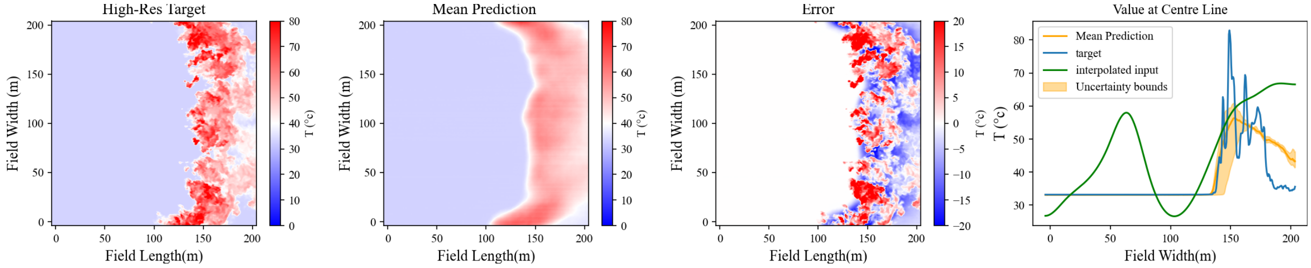
[70] Shuoyu Liu, Ying Luo, Liuliu Peng, Yan Jiang, Ercong Meng, and Bo Li. Wind pressure field reconstruction based on unbiased conditional kernel density estimation. *Journal of Wind Engineering and Industrial Aerodynamics*, 223:104947, 2022.



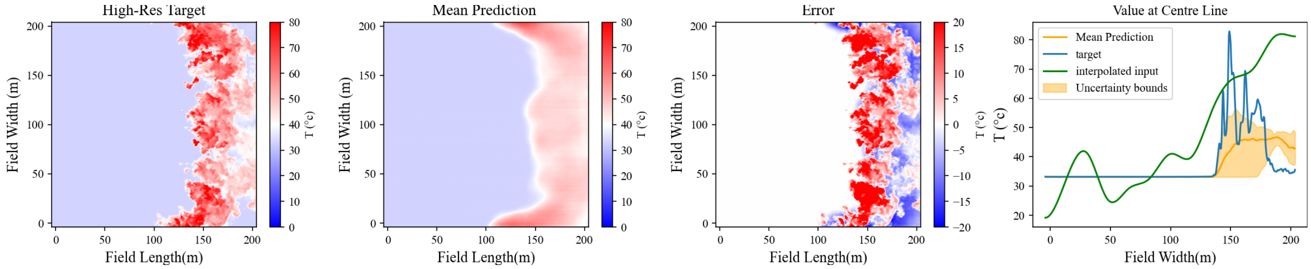
(a) Measurement noise= 0%



(b) Measurement noise= 10%



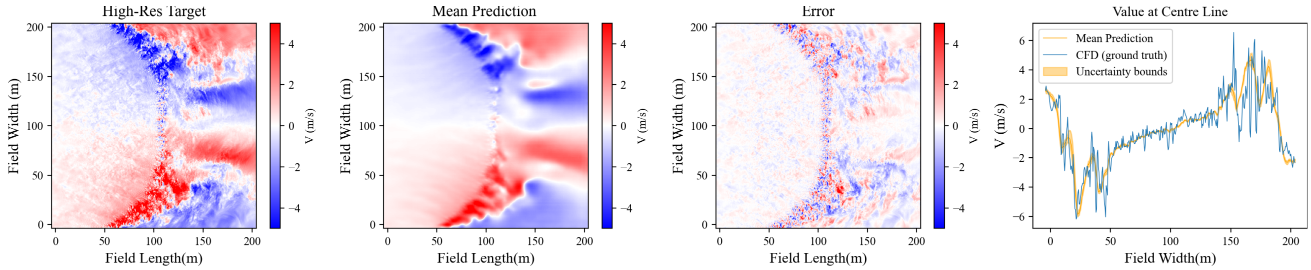
(c) Measurement noise= 30%



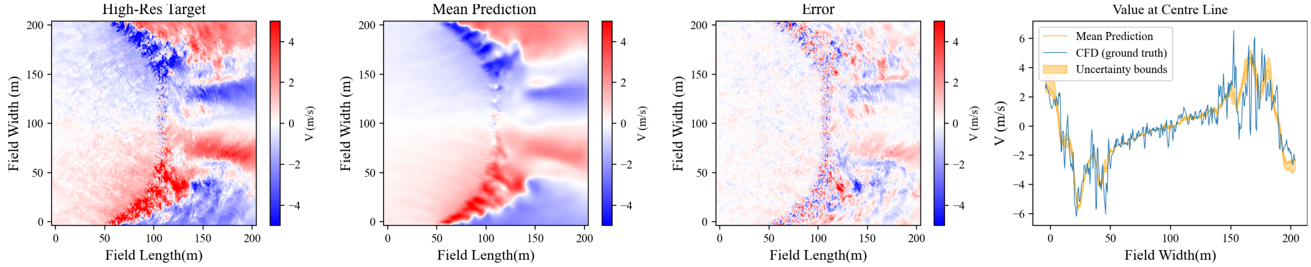
(d) Measurement noise= 50%

Figure S1.1: 49 (3) -T against different noise levels

AppendixS1. Extra Figures

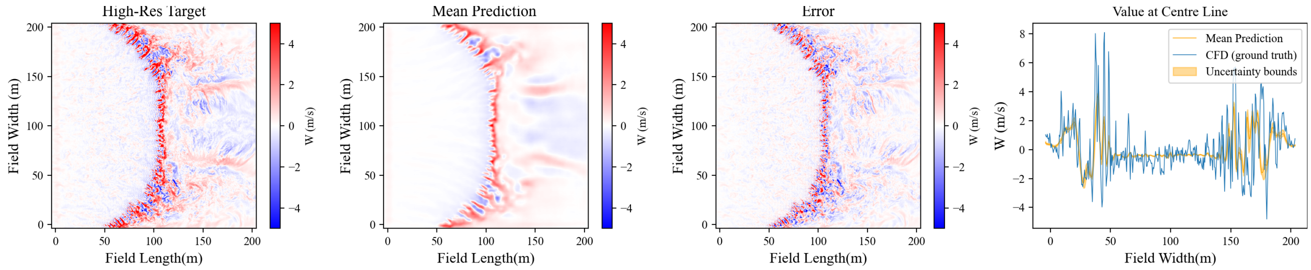


(a) Measurement noise= 0%

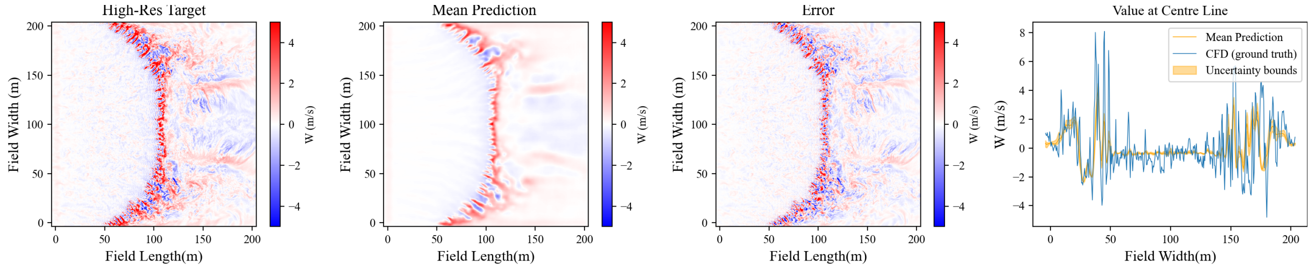


(b) Measurement noise= 50%

Figure S1.2: Results of the near-surface V-component of wind velocity estimation for case 49(3), $u_{10} = 4$ m/s, $H_g = 0.2$ m and $S = 0\%$, under different measurement noise levels, showing the high-resolution ground truth/target, the model's mean prediction, the reconstruction error, and a comparison of the estimated value along the centerline of the field with the ground truth (CFD)

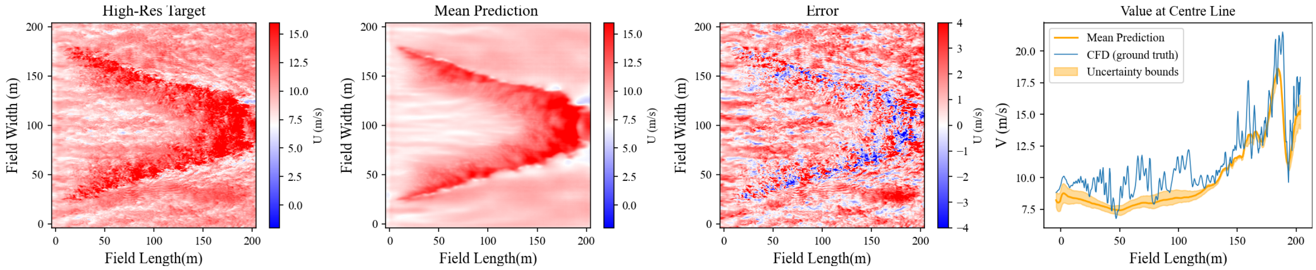


(a) Measurement noise= 0%

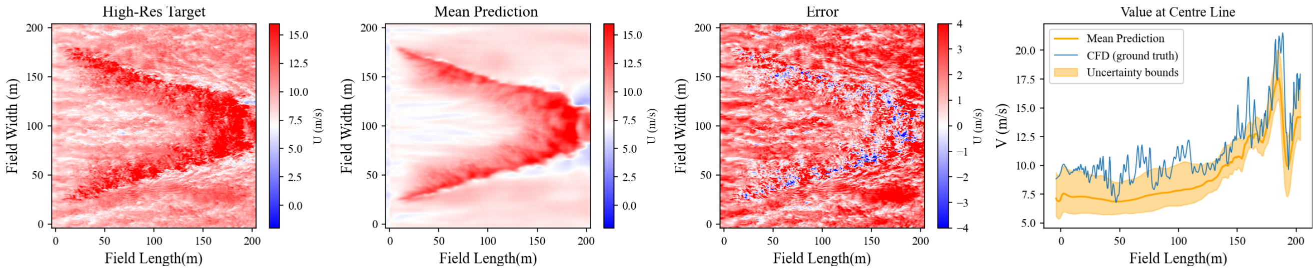


(b) Measurement noise= 50%

Figure S1.3: Results of the near-surface W-component of wind velocity estimation for case 49(3), $u_{10} = 4$ m/s, $H_g = 0.2$ m and $S = 0\%$, under different measurement noise levels, showing the high-resolution ground truth/target, the model's mean prediction, the reconstruction error, and a comparison of the estimated value along the centerline of the field with the ground truth (CFD)

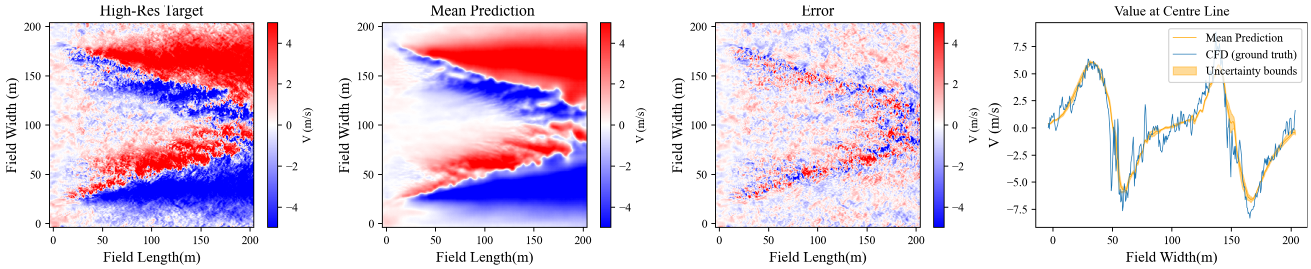


(a) Measurement noise= 0%

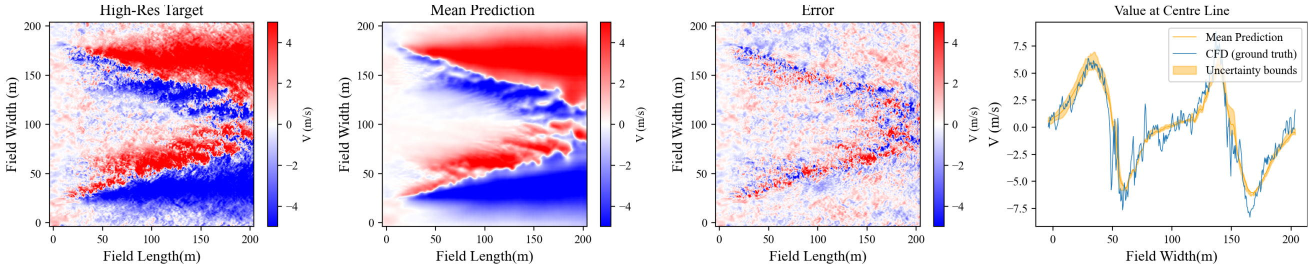


(b) Measurement noise= 50%

Figure S1.4: Results of the near-surface U-component of wind velocity estimation for case 49(3), u_{10} = 12 m/s, H_g = 0.5 m and S = +20%, under different measurement noise levels, showing the high-resolution ground truth/target, the model's mean prediction, the reconstruction error, and a comparison of the estimated value along the centerline of the field with the ground truth (CFD)

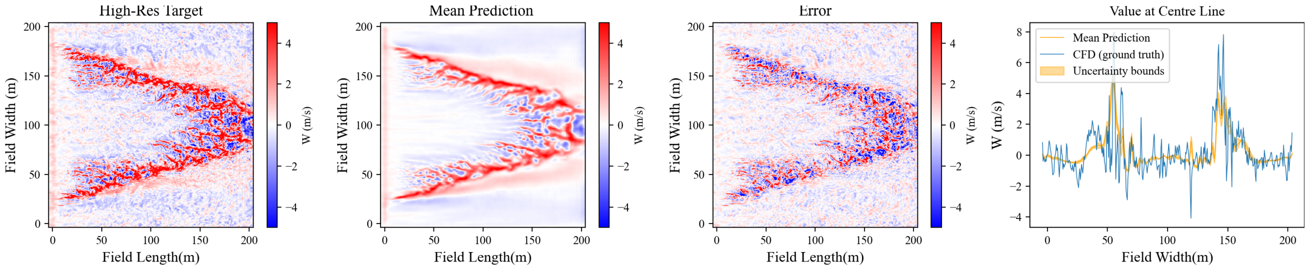


(a) Measurement noise= 0%

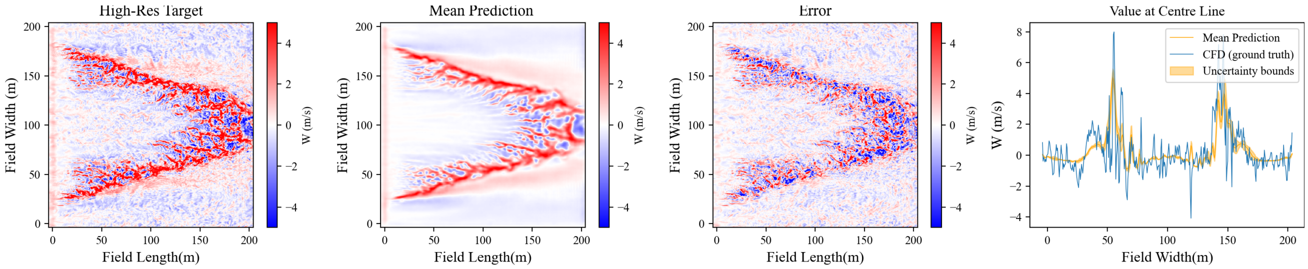


(b) Measurement noise= 50%

Figure S1.5: Results of the near-surface V-component of wind velocity estimation for case 49(3), u_{10} = 12 m/s, H_g = 0.5 m and S = +20%, under different measurement noise levels, showing the high-resolution ground truth/target, the model's mean prediction, the reconstruction error, and a comparison of the estimated value along the centerline of the field with the ground truth (CFD)

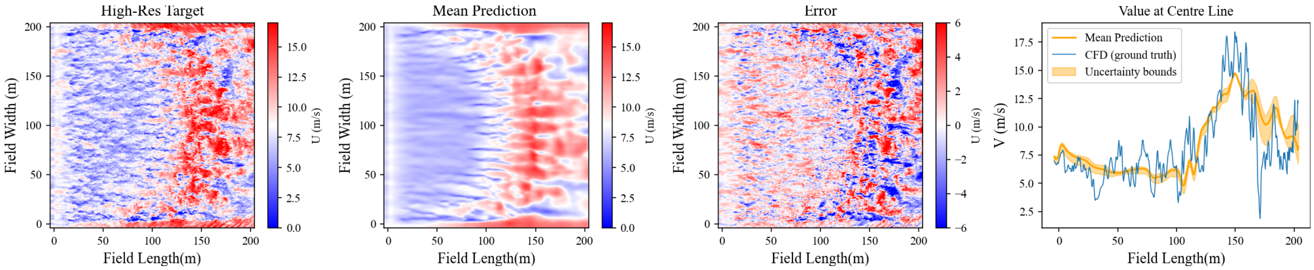


(a) Measurement noise= 0%

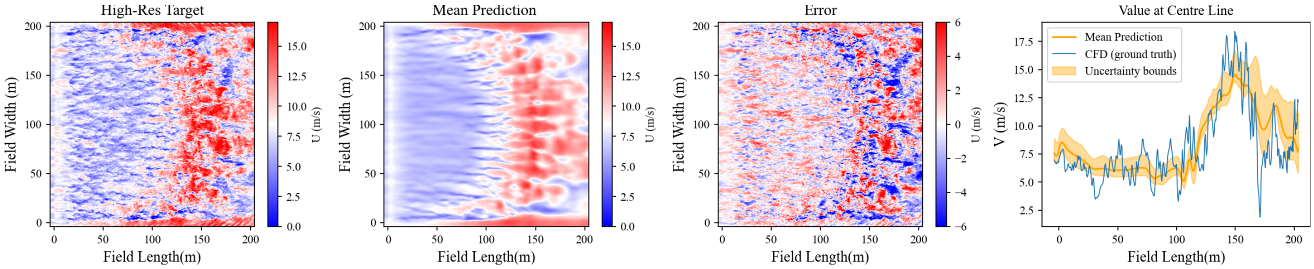


(b) Measurement noise= 50%

Figure S1.6: Results of the near-surface W-component of wind velocity estimation for case 49(3), u_{10} = 12 m/s, H_g = 0.5 m and S = +20%, under different measurement noise levels, showing the high-resolution ground truth/target, the model's mean prediction, the reconstruction error, and a comparison of the estimated value along the centerline of the field with the ground truth (CFD)

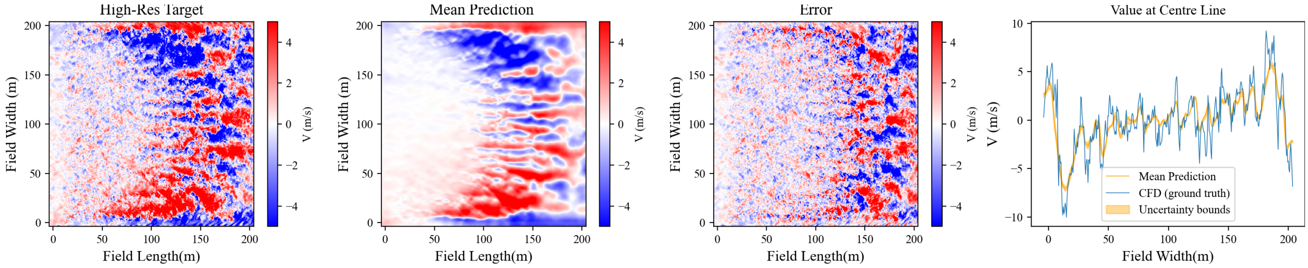


(a) Measurement noise= 0%

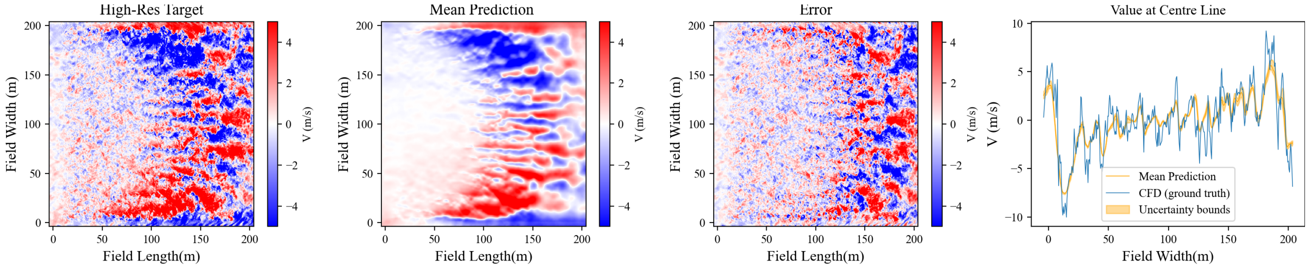


(b) Measurement noise= 50%

Figure S1.7: Results of the near-surface U-component of wind velocity estimation for case 49(3), u_{10} = 12 m/s, H_g = 1 m and S = −20%, under different measurement noise levels, showing the high-resolution ground truth/target, the model's mean prediction, the reconstruction error, and a comparison of the estimated value along the centerline of the field with the ground truth (CFD)

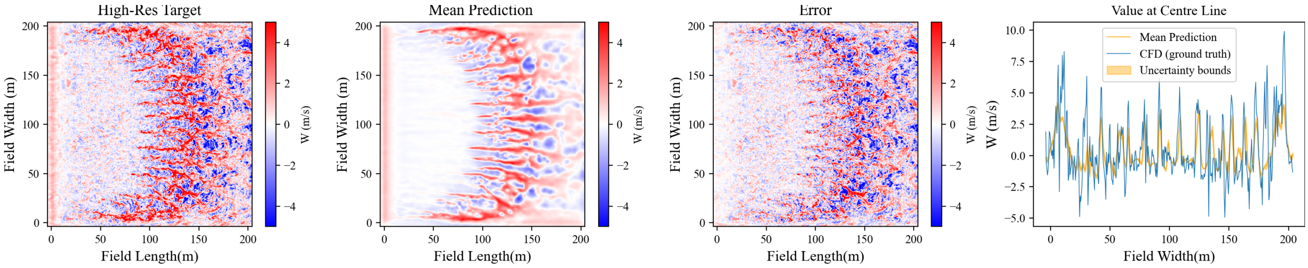


(a) Measurement noise= 0%

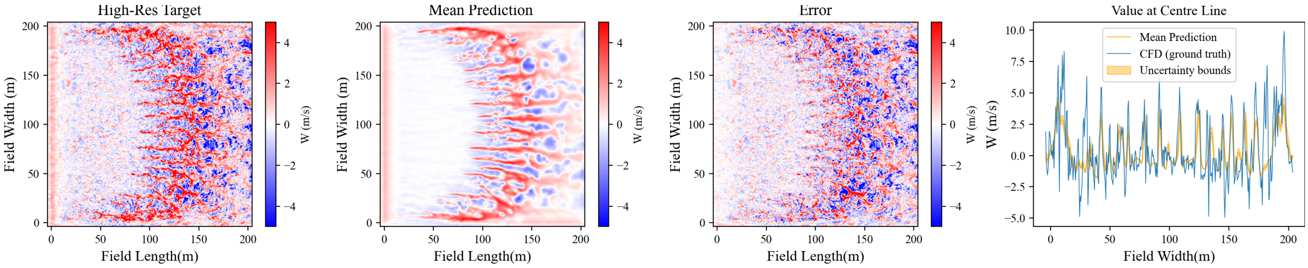


(b) Measurement noise= 50%

Figure S1.8: Results of the near-surface V-component of wind velocity estimation for case 49(3), $u_{10} = 12$ m/s, $H_g = 1$ m and $S = -20\%$, under different measurement noise levels, showing the high-resolution ground truth/target, the model's mean prediction, the reconstruction error, and a comparison of the estimated value along the centerline of the field with the ground truth (CFD)



(a) Measurement noise= 0%



(b) Measurement noise= 50%

Figure S1.9: Results of the near-surface W-component of wind velocity estimation for case 49(3), $u_{10} = 12$ m/s, $H_g = 1$ m and $S = -20\%$, under different measurement noise levels, showing the high-resolution ground truth/target, the model's mean prediction, the reconstruction error, and a comparison of the estimated value along the centerline of the field with the ground truth (CFD)