Wind Energy

WILEY

RESEARCH ARTICLE OPEN ACCESS

An Engineering Model for Static Yawed Wind Turbines Based on Actuator Line Simulations and Symbolic Regression

Haoyuan Sun | Andrea Sciacchitano | Wei Yu

Faculty of Aerospace Engineering, Delft University of Technology, Delft, the Netherlands

Correspondence: Haoyuan Sun (H.Y.Sun@tudelft.nl)

Received: 16 December 2025 | **Revised:** 10 March 2026 | **Accepted:** 24 March 2026

Keywords: induction | rotor aerodynamics | symbolic regression | yaw engineering model | yaw misalignment

ABSTRACT

Yaw engineering models are commonly used as add-ons to the industrial blade element momentum (BEM) framework to improve load and power predictions by accounting for the skewed wake effect. However, existing yaw engineering models show noticeable limitations in accurately predicting the induced velocity distribution across the blade span. In this study, we employ a genetic symbolic regression （SR） approach to develop a new set of yaw engineering models for both the normal and tangential induced velocities of a static yawed wind turbine. The model regression is performed using simulation data from Reynolds-averaged Navier–Stokes (RANS) simulations with an actuator line model (ALM) of the NREL 5-MW wind turbine, covering a range of yaw angles (γ) and thrust coefficients (C_T) over which the skewed wake effect is dominant. The regressed models are selected based on an optimal trade-off between accuracy and complexity, with complexity constrained to remain comparable with Branlard's yaw engineering model. The selected models are subsequently verified using three unseen cases that span different operating conditions and wind turbine models. Verification is performed through a series of evaluations, including generalization performance tests, implementation within the BEM framework to assess their aerodynamic performances, and quantitative errors and loading analyses. The results demonstrate that the proposed models improve both the amplitude accuracy and azimuthal phase of induced velocities compared with the existing models of Coleman and Branlard, enabling it to accurately capture the phase of the peak aerodynamic forces across each annulus and to predict the nonrestoring yaw moment occurring in the inboard region of the turbine, which other models fail to reproduce.

1 | Introduction

Blade element momentum (BEM) theory [1] is widely regarded as the most effective approach for balancing computational efficiency and accuracy in wind turbine simulations, particularly for large-scale industrial applications. This effectiveness arises from its foundational assumption of an infinite number of blades, along with the integration of supplementary engineering models that account for complex aerodynamic phenomena typically captured only by higher fidelity methods such as vortex models or computational fluid dynamics (CFD). These phenomena include blade root and tip losses, high-induction breakdown, dynamic stall, dynamic inflow, and skewed wake effects, among others [2, 3]. Yaw engineering models, in particular, are used to account for the nonuniform velocity distribution across the rotor disk caused by the skewed wake behind a yawed wind turbine [4, 5]. When a wind turbine operates in yaw, the wake becomes skewed, and vortices shed from the blade tips and root follow asymmetric trajectories downstream [6, 7]. This results in nonuniform induced velocities across the rotor plane, leading to nonuniform aerodynamic loading [8], that the standard BEM framework,

which relies on uniform loading assumptions, cannot inherently capture without the inclusion of appropriate engineering models [1].To date, two primary types of yaw engineering models have been developed to account for the skewed wake effect: (1) vortex-based models and (2) data-driven models. The vortex-based models traces back to 1926, when Glauert laid the foundation of autogyro theory and provided a simple estimation of the induced velocity for small angles of incidence on helicopter rotors [9]. Building on this work, Coleman proposed in 1945 a simplified cylindrical vortex wake model [10] representing the wake behind a helicopter rotor to derive a formula for the normal induced velocity on the rotor disk. This model has since been adapted for use in wind turbine simulations under yawed conditions, as expressed in Equation (1). Unlike classical BEM, which assumes uniform induced velocity along each radial annulus, Coleman's model applies an analytical nonuniform correction to the disk averaged induced velocity $\overline{V_i}$ to account for the asymmetric velocity distribution over a yawed rotor disk. The model introduces dependencies on three key variables: the normalized radial position r/R, the azimuthal angle θ, and the skewed wake angle χ (the angle between the rotor axis and the wake centerline). However, Coleman's model is based on the assumptions of an infinite tip-speed ratio and constant rotor circulation. Under these conditions, the bound vorticity, as well as the longitudinal tip vorticity and root vortex strength, approaches zero. Therefore, only the tangential trailed tip vortices remain significant and are considered in the model. In addition, the model assumes an infinite number of blades and a rigid cylindrical wake by ignoring the wake expansion. These assumptions limit the model's applicability under realistic operating conditions.

$$V_i = \overline{V_i}\left(1 + \frac{r}{R}\tan\frac{\chi}{2}\sin\theta\right) \tag{1}$$

$$F_r = \frac{r}{R} + 0.4\left(\frac{r}{R}\right)^3 + 0.4\left(\frac{r}{R}\right)^5 \tag{2}$$

Building on Coleman's model, Øye extended the formulation by replacing the linear radial dependency (r/R) with a polynomial function F_r that accounts for wake expansion, as expressed in Equation (2). The radial variation term F_r was determined through a curve-fitting procedure applied to results from a vortex ring model [11]. Within the same cylindrical vortex wake framework applied by Coleman, Branlard extended the approach by incorporating a more complete set of vortex components [12, 13] and subsequently developed a new yaw engineering model based on this enhanced framework [14]. Unlike Coleman's model, which considers only the tangential tip vortex, Branlard's model accounts for both tangential and longitudinal tip vortices, along with bound and root vortices, showing improved performance near the blade midspan. By explicitly modeling the complete vortex system, Branlard's model eliminates Coleman's restrictive infinite tip-speed ratio assumption and also enables the prediction of the tangential induced velocity, which was previously either neglected [10] or assumed to behave identically to the normal component [1]. The complete formulation of Branlard's model is provided in Appendix A. Branlard further systematically evaluated the three remaining assumptions underlying the cylindrical vortex wake model [15]. First, he investigated the impact of wake distortion using a vortex code by comparing cases with and without wake distortion. The results indicate that the yaw engineering model, which relies on the rigid wake assumption, remains valid, as the nondistorted wake case captures the dominant azimuthal and radial trends, with only minor phase and amplitude deviations relative to the distorted-wake case. Nevertheless, because the thrust coefficient of the considered cases was not reported, it is not certain whether this finding extends to higher loading conditions, where wake distortion may have a more significant impact. The second key assumption is an infinite number of blades and can be relaxed by applying Glauert's tip-loss correction. The only remaining idealization assumption is the uniform rotor circulation. Under this assumption, both Coleman's and Branlard's models exhibit a misprediction in the phase of normal induced velocities near the blade midspan [15]. Branlard addressed this limitation by demonstrating that using a superposition of concentric vortex cylinders, rather than a single cylinder with uniform circulation, can improve the phase alignment of induced velocities and the resulting aerodynamic forces. While this approach improves physical fidelity, it also increases the complexity of the BEM framework, as it requires radial discretization of vortex cylinders from blade root to tip [16], thereby complicating the implementation.

In addition to vortex-based models, data-driven models have been developed by fitting high-fidelity numerical and experimental results using traditional mathematical techniques. To overcome the limitation of Coleman's model, Schepers developed a more advanced model based on wind tunnel measurements [17], explicitly accounting for root vortex effects. The experimentally obtained normal induced velocity was approximated by a second-order Fourier series in the azimuthal angle, with the Fourier amplitudes and phases parameterized as functions of radial position and yaw angle using a least-squares regression approach, as expressed in Equation (3). This model was subsequently verified and refined using various wind turbine measurement datasets, achieving a considerable improvement in load prediction. However, a later study demonstrated that Schepers' model exhibits phase misalignment over the azimuthal direction at the blade tip region when applied to full-scale wind turbines [18]. This discrepancy may result from the model's overprediction of the root vortex because it is based on experimental data from a small turbine with a 1.2 m rotor diameter that generates a strong root vortex [19]. As a further development, Rahimi applied the same second-order Fourier fitting to data from actuator line simulations of three multi-MW class wind turbines and introduced tip-speed ratio as an additional dependency of the Fourier coefficients which was not covered in Schepers' model. The resulting model shows improved accuracy, outperforming both the conventional Glauert's model and Schepers' model.

$$V_i = \overline{V_i}\left[1 - A_{1,i}\cos(\theta - \psi_1) - A_{2,i}\cos(\theta - \psi_2)\right] \tag{3}$$

Currently, the most accurate and accessible engineering-level yaw modeling approaches, to the best of the authors' knowledge, involve embedding Branlard's model within a superposition of concentric vortex cylinders, which increases the complexity of BEM implementation. On the other hand, in Schepers' and Rahimi's data-driven models, the Fourier coefficients were obtained separately for each inflow condition and radial position and have not been fully released to the

public, leaving a gap in the availability of a reliable and user-friendly engineering model for yawed wind turbine simulations. Meanwhile, further improvements in yaw engineering models appear to have reached a bottleneck. From a physics-based perspective, Branlard's model already incorporates comprehensive vortex components and has validated all underlying assumptions. From a data-driven perspective, classical mathematical fitting is a manual and time-consuming process sensitive to the choice of function shapes and can easily converge to local optima, thereby posing practical challenges for model development. To address this challenge, we introduce symbolic regression (SR) here as an efficient and automated methodology for model discovery.

SR is a machine learning technique that aims to discover mathematical expressions that best describe a given dataset [20–22]. Unlike traditional regression methods that fit parameters within a fixed model structure, SR explores a broad space of mathematical expressions to identify models that are both accurate and interpretable. A distinguishing feature of SR is its highly customizable search space: Users can guide the search process by imposing constraints through the specification of allowed mathematical operators and input variables (features), effectively tailoring the model discovery process to the problem at hand [23]. This flexibility enables SR to leverage domain knowledge, such as physical constraints or symbolic forms, making it well suited for physics-informed modeling, which present a novel approach to modeling. SR has gained widespread application across various scientific fields [24, 25] but remains underutilized in wind energy research. Wang et al. employed SR in combination with a double-Gaussian equation structure to derive a compact and physically grounded wake model with strong predictive performance [26]. Valsaraj et al. developed a SR method for wind speed extrapolation that outperforms traditional power law approaches, enabling accurate wind resource assessment using shorter measurement masts [27]. Wang et al. [28] introduced a SR-enhanced dynamic wake meandering (DWM) model that embeds data-derived symbolic expressions into the DWM framework to reconstruct volumetric forcing and boundary conditions. By achieving equation-level closure through interpretable equations extracted from large eddy simulation data, their approach significantly improves near-wake physical consistency while preserving the computational efficiency. However, SR has not yet been applied to the development of yaw engineering models, and its effectiveness in this context remains unexplored, which motivates the present study.

In this study, SR is employed to develop a new set of yaw engineering models for both the normal and tangential induced velocities of a static yawed wind turbine. The input features used in the SR process are informed by Branlard's model, which embeds comprehensive physical insights to ensure consistency with first principles while enabling data-driven optimization. The model is regressed based on data from Reynolds-averaged Navier–Stokes (RANS) simulations coupled with the actuator line model (ALM) [29] and verified using three independent unseen sets. The SR model is subsequently implemented within an in-house BEM framework to evaluate model performance, quantify errors, and conduct loading analysis, alongside Coleman's and Branlard's models. The primary objective of this study is to propose an improved, compact, and user-friendly model that can be readily integrated into existing BEM codes. To the authors' knowledge, this work represents the first application of SR to the development of yaw engineering models.

2 | Methodology

This section presents the methodologies adopted in the study, organized into four subsections. Section 2.1 outlines the simulation setup, training datasets, coordinate convention, and extraction of induced velocity from ALM simulations. Section 2.2 introduces the SR algorithm, including the selection of features and target variables, the operators applied, and the constraints imposed during model regression. Section 2.3 describes the model selection procedure, while Section 2.4 details the procedure used to verify the selected model.

2.1 | CFD Simulation Setup

To generate input data for SR, a series of simulations has been carried out on a standalone, three-bladed horizontal-axis wind turbine subjected to a range of yaw misalignment angles, encompassing both positive and negative values. In the present study, the wind turbine is represented using the ALM, which introduces certain simplifications compared with fully blade-resolved simulations or real turbine systems as indicated by Rahimi et al. [18]. While the ALM approach captures the main rotor aerodynamic effects, some detailed blade–flow interactions may not be fully resolved. As a result, slight differences may arise compared with more complex simulations or experimental measurements. Nevertheless, the ALM method is widely adopted in wind turbine studies due to its good balance between computational efficiency and physical accuracy. This study employs the NREL 5-MW reference turbine (Jonkman et al. [30]) as the rotor model, owing to its extensive validation and widespread use in the literature [31, 32]. Key specifications include a rotor diameter of $D = 126\,\mathrm{m}$, a below-rated wind speed of $V_\infty = 8.0\,\mathrm{m/s}$, and a tip-speed ratio of $\lambda = 7.55$. Unless otherwise stated, all simulations are conducted under this condition. To isolate the aerodynamic effects of yaw misalignment, the model excludes structural and environmental complexities such as tower effects, tilt, coning, ground interaction, wind shear, control systems, inflow turbulence, and aeroelastic behavior.

The simulations are performed using the open-source CFD framework *OpenFOAM v2306* (OpenCFD Ltd. [33]), together with *turbinesFoam*, an ALM module originally developed by Bachant et al. [34]. The flow is treated as incompressible and Newtonian, with constant density $\rho = 1.225\,\mathrm{kg/m^3}$ and kinematic viscosity $\nu = 1.5 \times 10^{-5}\,\mathrm{m^2/s}$. Thermal effects and Coriolis accelerations are neglected to focus on the fundamental flow mechanisms. The governing equations are solved using the RANS framework, with turbulence closure provided by the shear stress transport (SST) variant of the k–ω model [35], selected for its demonstrated reliability and robustness in wind energy research [36, 37]. Second-order central differencing is used for spatial discretization, while temporal integration is performed using a blended Crank–Nicolson

scheme [38] with a weighting factor of 0.9, providing a balance between numerical stability and second-order temporal accuracy. Pressure-velocity coupling is handled using the PIMPLE algorithm [33], which combines PISO [39] and SIMPLE [40] algorithms to maintain stability and facilitate convergence in unsteady RANS simulations. The computational domain has a uniform inflow velocity at the inlet, a zero-gradient condition at the outlet, and slip boundary conditions along the lateral, top, and bottom surfaces.

Turbulent inflow conditions are imposed using the divergence-free synthetic eddy method (DFSEM) [41]. This method generates synthetic turbulence with prescribed properties, such as turbulence intensity, length scales, and anisotropy, enabling realistic turbulent inflow conditions to be introduced into the computational domain. Compared with precursor simulation approaches, DFSEM significantly reduces the computational cost while still providing representative turbulence structures at the inlet. To generate a nearly laminar inflow condition while avoiding numerical divergence, the inflow turbulence intensity is set close to zero to mimic idealized conditions [42]. The turbulent kinetic energy at inlet is specified as $k = 2\times10^{-4}\ \mathrm{m^2/s^2}$, corresponding to a turbulence intensity of

$$\mathrm{TI} = \frac{\sqrt{2k/3}}{V_\infty} \approx 0.045\,\% \tag{4}$$

where V_∞ is the free-stream flow velocity. The turbulence-specific dissipation rate at inlet is set to $\omega = 2\times10^{-4}\ \mathrm{s^{-1}}$, based on the relation

$$\omega = \frac{\sqrt{k}}{C_\mu^{0.25} L} \tag{5}$$

with $C_\mu = 0.09$ as an empirical constant and L denoting the reference length scale. These settings are chosen to provide a clean and controlled inflow environment, allowing for the investigation of fundamental aerodynamic behaviors without the influence of upstream disturbances.

$$\boldsymbol{f}_{2D}(r) = (\boldsymbol{L},\boldsymbol{D}) = \frac{1}{2}\rho V_{\mathrm{rel}}^2 c\left(c_l(\alpha)\hat{\boldsymbol{e}}_L + c_d(\alpha)\hat{\boldsymbol{e}}_D\right) = f_n\hat{\boldsymbol{e}}_n + f_t\hat{\boldsymbol{e}}_t, \tag{6}$$

$$V_{\mathrm{rel}} = \sqrt{V_n^2 + V_t^2}, \quad \phi = \arctan\left(\frac{V_n}{V_t}\right) = \alpha + \beta, \tag{7}$$

$$\boldsymbol{f}_{\mathrm{body}}(\boldsymbol{x}) = \sum_{i=1}^{B}\int_0^R f_{\mathrm{tip}}(r_i)\boldsymbol{f}_{2D}(r_i)\,\eta_\epsilon\left(\left\|\boldsymbol{x} - (r_i\hat{\boldsymbol{e}}_i + \boldsymbol{p}_R)\right\|\right)dr_i, \tag{8}$$

$$\eta_\epsilon(d) = \frac{1}{\epsilon^3\pi^{3/2}}\exp\left[-\left(\frac{d}{\epsilon}\right)^2\right], \quad f_{\mathrm{tip}}(r) = \frac{2}{\pi}\arccos\left[\exp\left(-\frac{B(R-r)}{2r\sin\phi}\right)\right] \tag{9}$$

The computational domain in this study is based on the configuration presented by Li et al. [43] and has been validated through a grid independence test; therefore, it is not reproduced. A Cartesian coordinate system is adopted, where the x axis aligns with the inflow direction, the y axis extends horizontally perpendicular to the flow, and the z axis points vertically upward as shown in Figure 1. The computational domain measures $12.5D$ in the streamwise direction and $5D$ in both the lateral and vertical directions, resulting in overall dimensions of $12.5D \times 5D \times 5D$. Figure 1 displays cross-sectional views of the computational mesh at $x/D = 0$ (left) and $y/D = 0$ (right), illustrating the hierarchical mesh refinement strategy. The term "Level" indicates the refinement tier, with Level 4 corresponding to a finest grid resolution of $\Delta = D/80$, which is comparable with the actuator point spacing [44]. Each decrement in refinement level doubles the local grid spacing, and the resulting computational mesh consists of approximately 10.48 million cells. The turbine rotor is centered at the origin and rotates clockwise when viewed from upstream. For simulations of the NREL 5-MW turbine operating at below-rated conditions, each full rotor revolution is discretized into 360 time steps, yielding a temporal resolution of $\Delta t = 0.0182\mathrm{s}$. This setup ensures that the rotor tip advances less than one grid cell per time step, typically staying below 0.7Δ, which provides adequate temporal resolution [44].

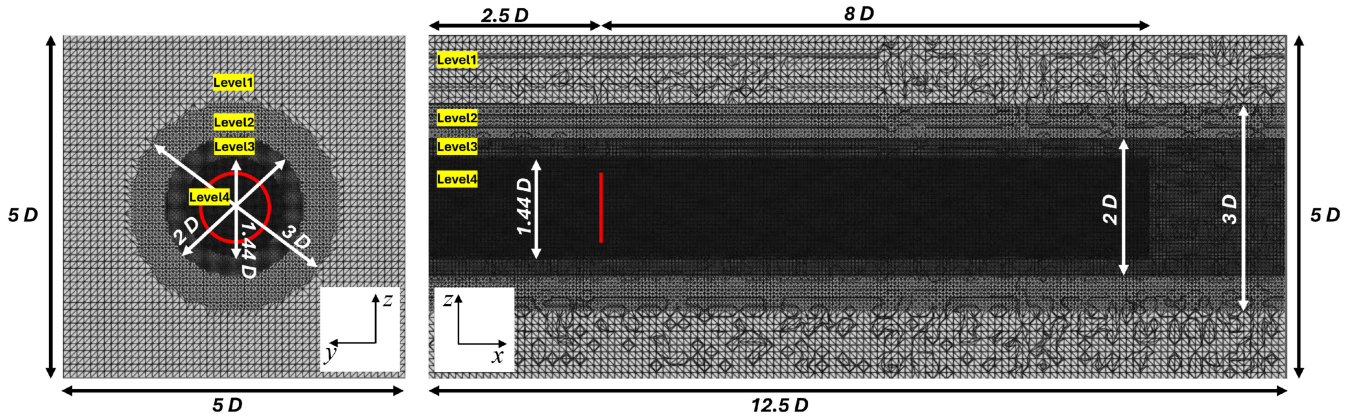

FIGURE 1 | Mesh layout used for ALM simulations of the NREL 5-MW turbine with the turbine position marked in red. Left: Front view (at $x/D = 0$) showing concentric refinement zones centered on the rotor disk. Right: Side view (at $y/D = 0$) illustrating vertical grid refinement and the extent of the high-resolution region, which spans $1.44D$ in height at the finest mesh and spans $2D$ and $3D$ for the two coarser levels, respectively. Four levels of grid refinement are applied, with each level doubling the grid spacing relative to the previous one.

In the ALM, each turbine blade is represented by 40 evenly distributed points along the spanwise direction. To capture the influence of the hub region, an additional point is positioned at the rotor center and modeled as a bluff body [45]. This central element applies a drag force characterized by a coefficient $C_d = 0.3$, with the reference area taken as πr_{hub}^2, where $r_{\text{hub}} = 1.5$ m denotes the hub radius [43]. Aerodynamic forces acting on the actuator points are evaluated using BEM theory [1], where the local lift (L) and drag (D) components are computed based on the local relative velocities (V_{rel}) and pretabulated airfoil data. These force vectors $\boldsymbol{f}$ are then smoothly projected onto the background mesh via a Gaussian kernel function η_ϵ [29], following the formulation in Equations (6–9). This regularization ensures numerical stability and maintains physical consistency by avoiding force discontinuities in the flow field. The equations involve the rotor radius R, the radial distance r along the spanwise direction of the blade, the local chord length c, the blade twist angle β, and the number of blades B. To smoothly reduce aerodynamic loads near the blade root and tip, a tip-loss correction factor f_{tip} based on the Glauert model is applied [46]. Following established ALM practices [47, 48], the smoothing parameter ϵ in the regularization kernel is assigned a value equal to twice the local mesh size near the rotor, corresponding to 2Δ at Level 4 (see Figure 1).

Based on the setup described above, and to improve the generalization capability of the models developed in this study, additional turbine operating states with varying thrust coefficients (C_T) are considered by intentionally pitching the blades. In total, three pitch settings of –2°, 0°, and 2° are simulated at the same inflow velocity and tip-speed ratio as summarized in Table 1. The resulting rotor-averaged C_T values are shown in Figure 2, ranging from 0.65 to 0.86 for the nonyawed case. Because the skewed wake effect becomes negligible at low tip-speed ratios due to the increasing influence of the advancing and retreating effect [17], above-rated wind speeds are not considered in this study. Therefore, this C_T range is deemed sufficient to represent C_T variation of typical turbines operating near the optimal tip-speed ratio [49, 50], prior to the onset of active blade pitching. The C_t at each blade element is computed as

$$C_t = \frac{T}{\frac{1}{2}\rho(V_\infty \cos\gamma)^2 A} \tag{10}$$

where γ is the yaw angle and A is the rotor swept aera. For each training (Tr) case in Table 1, a total of seven simulations are performed with yaw angles ranging from –30° to +30° in 10° increments. In each simulation, the last 20 revolutions are used and averaged over all azimuthal positions swept by a single blade during one full revolution [43]. Each simulation produces velocity fields for 40 blade elements at 360 azimuthal positions, resulting in 100,800 velocity data points across all simulations for each case.

For clarity, the coordinate convention is defined here and consistently applied throughout this paper. Following the right-hand rule, the yaw angle is defined as positive for counterclockwise rotation when viewed from above (see Figure 3, left) and negative for clockwise rotation. Because the wind turbine rotates clockwise in the simulations (viewed from upstream), the azimuthal angle θ is defined as positive in the clockwise direction, with $\theta = 0°$ corresponding to the blade pointing vertically upward (see Figure 3, right). For the rotor plane, the local coordinate system (x', y', z') is defined to coincide with the computational domain coordinate system when the turbine is not yawed, with the normal direction (x' axis) aligned with the streamwise flow. When the blade is at the 12 o'clock position, the tangential direction (y' axis) along the blade points leftward opposite to the rotating direction, and the radial direction (z' axis) extends outward from the hub along the blade span. The rotor plane coordinate system rotates following the yaw angle convention shown in Figure 3.

The velocity at each actuator line element is sampled using the line-average method as recommended [51]. The axial and tangential velocities V_n, V_t at each element are then computed and transformed into the corresponding induced velocities V_{ni}, V_{ti}, which serve as the input for SR. Figure 4 illustrates the velocity triangle for a blade element (actuator line point) on a positive yawed wind turbine, where x, y and x', y' correspond to the computational domain and airfoil-local coordinates, respectively. Note that for a yawed rotor, the inflow does not impinge on the rotor plane perpendicularly. This reduces the effective normal velocity component from the free-stream velocity V_∞ to $V_\infty \cos\gamma$ and introduces an additional tangential component $V_\infty \sin\gamma \cos\theta$, resulting in azimuthally varying in-plane velocity

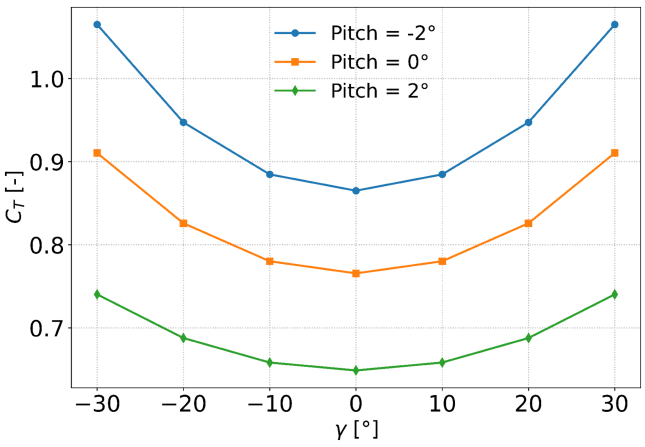

FIGURE 2 | Variation of rotor-averaged thrust coefficient (C_T) with yaw angle for different blade pitch settings (–2°, 0°, and 2°) at the inflow velocity of $V_\infty = 8.0$ m/s and the tip-speed ratio of $\lambda = 7.55$.

TABLE 1 | Wind turbine operating parameters of the training (Tr) cases used for SR.

Case	Turbine model	Wind speed (m/s)	Tip-speed ratio (-)	Yaw angle (°)	Pitch (°)
Tr.1	NREL-5 MW	8.0	7.55	± 30, Δ10	0
Tr.2	NREL-5 MW	8.0	7.55	± 30, Δ10	2
Tr.3	NREL-5 MW	8.0	7.55	± 30, Δ10	–2

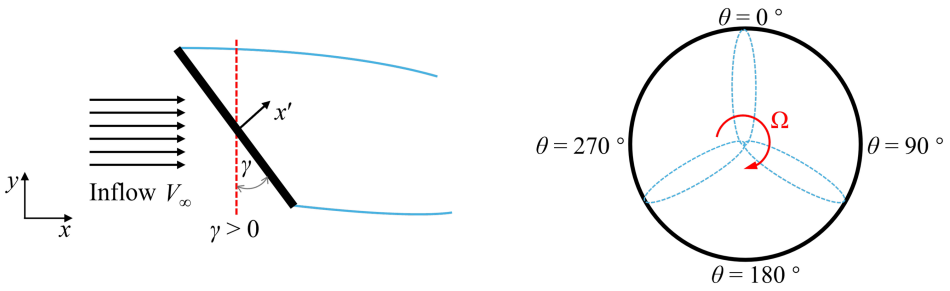

FIGURE 3 | Coordinate conventions for yaw and azimuthal angles. The left panel (top view) illustrates the yaw angle γ, defined as positive when the rotor yaws to the left, along with the inflow velocity V_∞. The right panel (viewed from upstream) illustrates the azimuthal angle θ, defined relative to the rotor's rotation direction: $\theta = 0°$ is at the top (12 o'clock) position and increases clockwise, while the rotor rotates in the direction indicated by the angular velocity Ω.

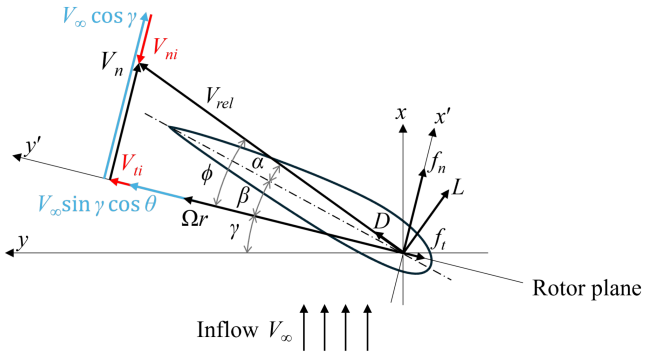

FIGURE 4 | Velocity triangle for a blade element on a yawed wind turbine ($\gamma > 0$), illustrating the induced velocity in red and the velocity component caused by yaw misalignment in blue when the blade is at 0° azimuth (vertical position). The local twist angle β includes both the blade pretwist and the applied static pitch angle.

and, consequently, a varying angle of attack (α), commonly referred to as the advancing and retreating effect [7]. The normal and tangential induced velocity V_{ni} and V_{ti} are then computed as follows. Verification of the induced velocities and aerodynamic forces obtained from the ALM setup is provided in Appendix B.

$$V_n = V_x \cos\gamma + V_y \sin\gamma, \quad V_{ni} = V_\infty \cos\gamma - V_n, \tag{11}$$

$$V_t = (V_x \sin\gamma - V_y \cos\gamma)\cos\theta - V_z \sin\theta, \quad V_{ti} = V_t - V_\infty \sin\gamma \cos\theta - \Omega r. \tag{12}$$

2.2 | Symbolic Regression

In this study, the open-source package PySR [23] is used to perform SR and fit analytical expressions for the nonuniform induced velocity of a static yawed wind turbine. PySR is chosen as it was specifically developed for scientific discovery, offering advantages such as generating interpretable equations, efficient equation search through just-in-time compilation, and high customizability with support for user-defined operators, loss functions, and constraints to guide the search toward physically meaningful solutions. A detailed explanation of the analytical expression fitting procedure is provided below. Unless otherwise specified, the setups for the normal and tangential induced velocities are treated identically in the subsequent regression.

The first step is to identify the features used in the final SR model (i.e., the input variables X, as opposed to the target variable Y). PySR's flexible customization capabilities allows arbitrary features to be defined, enabling the integration of physical insights directly into the regression process. The selection of features is informed by observations from Branlard's model [15] (see Appendix A), which, as discussed in Section 1, incorporates the full vortex system to capture the main physical features. Based on this understanding, eight key features are selected, which are expected to be the most relevant to the target model:

$$X \in \left\{\tilde{r},\ \overline{C_t},\ \sin\theta,\ \cos\theta,\ \tan\gamma,\ \tan(\gamma/2),\ \sin\gamma,\ \cos\gamma\right\}$$

where $\tilde{r} = r/R$ is the radial position normalized by the rotor radius. Yaw angle (γ) is used, following the convention adopted in Schepers' model [17]. Features such as $\sin 2\theta$, which arises from the longitudinal tip vortex, are excluded to simplify the regression, as Branlard [14] pointed out that the influence of longitudinal tip vorticity is relatively minor compared with tangential tip and root vortices when the tip-speed ratio is substantially larger than one. The annulus-averaged $\overline{C_t}$ is included as a feature to represent the operating condition of the turbine and to capture the radial variation of the induced velocity. To prevent the occurrence of unphysical expressions during the regression, the azimuthal (θ) and yaw (γ) angles are not directly included as features, and trigonometric operators are not used directly either. Such usage may lead to expressions like $\sin(r/R)$, which are not physically interpretable and may misguide the search. Instead, precomputed features such as $\sin\theta$ are used to maintain consistency of physical units, with all features being dimensionless. The new yaw engineering models adopt the structure of classical Coleman's model, and the induced velocity is normalized by the azimuthally averaged induced velocity at each annulus. This formulation allows BEM implementation with independent annuli, as in Rahimi's work [19]. Accordingly, the regression target (Y) is defined as the induced velocity described in Section 2.1, normalized by the annulus-averaged value. This normalization results in a dimensionless target variable, thereby satisfying the requirement of dimensional consistency. The following operator set is employed in the regression:

$$+,\quad -,\quad \times,\quad \div,\quad \texttt{sqrt},\quad \texttt{square},\quad \texttt{cube}$$

To further guide the search and reduce redundancy in the expressions, several constraints are imposed. The operators

`square`, `cube`, and `sqrt` are restricted to act only on subexpressions with a maximum complexity (*Cplx*) of 5, while division is limited to subexpressions with a maximum *Cplx* of 15. Here, *Cplx* is defined as the total number of operators, constants, and features in an operator or expression. To prevent excessively nested exponentials, the `sqrt` operator may be nested with a `sqrt` or `cube`operator only once, and it cannot be nested inside a `square`, as this effectively reduces to an exponent of 1. Similarly, the `square` operator may be nested within another `square` only once, and the `cube` operator is limited to one nesting within either a `square` or another `cube`.

The default hyperparameters recommended by PySR are employed, as they have been optimized to provide a good balance between accuracy and computational efficiency across diverse benchmark problems [52–55]. The maximum allowable model *Cplx* during regression is determined following the guidance of Branlard's model [15], which exhibits *Cplx* levels of approximately 60–70 for both the normal and tangential induced velocities. It should be noted that, since the input features in this study have been preprocessed, the *Cplx* reported by SR does not represent the true model *Cplx*. Instead, it reflects a baseline that is already elevated compared with using the original unprocessed variables as features. As a result, the apparent *Cplx* observed during the SR search is lower than the actual *Cplx* of the final expression. Considering this and to avoid overly complex models, the maximum *Cplx* is set to 60 in this study, which ensures coverage of the range where meaningful expressions are most likely to emerge. Finally, SR, as a supervised machine learning technique, is applied to fit explicit and dimensionally consistent analytical expressions to the normalized induced velocity data of yawed wind turbines.

2.3 | Model Screening and Selection

The SR output from PySR provides a sequence of candidate expressions with increasing *Cplx*. To select the most appropriate model, a clear and systematic selection criterion is necessary. Each candidate is evaluated using two key metrics: the loss and *Cplx*. The loss, defined as the mean squared error (MSE) between model predictions and training data, quantifies accuracy, while *Cplx* reflects the structural intricacy of each expression, as introduced in Section 2.2. The model selection process proceeds as follows. All regressed candidate models are plotted with respect to model *Cplx* and logarithmic loss, forming a Pareto front [56]. To identify the most favorable trade-off between accuracy and *Cplx*, the fractional drop (f_d) is computed for each model, following the definition in Equation (13) [57]. This quantity reflects the relative gain in model accuracy per unit increase in *Cplx*. It is computed as the difference in logarithmic MSE, $\log(\mathrm{MSE}_i) - \log(\mathrm{MSE}_{i-1})$, divided by the corresponding change in complexity, $Cplx_i - Cplx_{i-1}$, between two adjacent models.

$$f_d{}^i = -\frac{\log(\mathrm{MSE}_i) - \log(\mathrm{MSE}_{i-1})}{Cplx_i - Cplx_{i-1}} \tag{13}$$

The model corresponding to the maximum f_d is selected as an initial candidate, as it offers the greatest improvement in accuracy relative to added complexity. However, as noted in prior work [26], high f_d values frequently arise in regions of very low *Cplx*, where the resulting expressions, despite showing steep relative gains, remain far from accurate and fail to capture the essential physics. These models are often underfitted and lack the capacity to represent the underlying relationships in the training data. To avoid selecting such inadequate models, a threshold set to 1.1 times the minimum loss across all regressed models is applied to restrict the candidate pool, ensuring that only models satisfying a baseline level of accuracy are considered.

2.4 | Model Verification Procedure

With the candidate expression obtained from SR, its generalization performance is first rigorously evaluated using unseen data. Table 2 summarizes all unseen (Ue) cases considered in this study. Initially, the SR model is assessed against reference ALM results under the same operating conditions as the Tr.1 case in Table 1, with intermediate yaw angles $\gamma \in [-25°,\ 25°]$ with a resolution of $\Delta\gamma = 10°$. This evaluation provides a quantitative, point-by-point comparison of the SR model predictions with the ALM data, measuring its consistency with the reference ALM results. Following this verification, the SR model is incorporated into an in-house BEM framework to evaluate its aerodynamic performance under operating conditions external to the training data. Two additional, distinct unseen cases are analyzed, as detailed in Table 2. The Ue.2 case simulates the rated operating condition of the NREL 5-MW turbine at $V_\infty = 11.4$ m/s and $\lambda = 7.0$, while the Ue.3 case investigates an entirely unseen turbine configuration, specifically the IEA 15-MW reference turbine [49], operating at $V_\infty = 9.0$ m/s and $\lambda = 9.0$. For both cases, the analysis is restricted to two representative yaw angles (5° and 25°) to effectively capture the most significant aerodynamic variations. The in-house BEM code employed in this study was developed following the methodology and detailed procedures described in [1], and its verification is provided in Appendix B.

A comparative analysis is first performed among three yaw engineering models, namely, Coleman's model, Branlard's model, and the SR model developed in this study, focusing on the prediction of normal and tangential induced velocities as well as the resulting aerodynamic forces. In addition to qualitative comparisons, a quantitative evaluation is conducted focusing on two key aspects. First, amplitude accuracy is assessed by calculating the root-mean-square error (RMSE) between velocities predicted by the BEM approach and those obtained

TABLE 2 | Wind turbine operating parameters of the unseen (Ue) cases for model verification.

Case	Turbine model	Wind speed (m/s)	Tip-speed ratio (-)	Yaw angle (°)	Pitch (°)
Ue.1	NREL 5-MW	8.0	7.55	±25, Δ10	0
Ue.2	NREL 5-MW	11.4	7.0	5, 25	0
Ue.3	IEA 15-MW	9.0	9.0	5, 25	0

from ALM simulations. Second, a significant drawback of current yaw engineering models is the inherent azimuthal phase misalignment observed in predicted induced velocities and aerodynamic forces, which is particularly pronounced in the inboard (below midspan) sections of the blade [18]. To investigate this, a phase alignment analysis is performed. This analysis uses cross-correlation [58] as the criteria to quantitatively evaluate the azimuthal synchronization between BEM predictions and ALM data for each radial annulus.

After completing qualitative and quantitative verification, the focus shifts to evaluating the impact of different yaw engineering models on loading predictions. A loading assessment [59] is performed by comparing the annulus-averaged yaw moments predicted by the SR model, as well as Coleman's and Branlard's models, against ALM results for yaw angles of $\gamma = 5°$ and 25°. The yaw moment (M_{yaw}) contributed by an individual actuator point is computed as

$$M_{\text{yaw}} = f_n \times r\cos(\theta + \pi/2) - f_t \cos\theta \times OH \tag{14}$$

OH denotes the overhang, which is 5 m for the NREL 5-MW turbine and 11.35 m for the IEA 15-MW turbine.

3 | Results and Discussion

This section presents the final regression models for the normal and tangential induced velocities, followed by model verification and performance analysis. In Section 3.1, the generalization performance and the selection process of the SR models are discussed. Next, in Section 3.2, the selected SR models are implemented within the BEM framework. The SR-improved models are compared with those proposed by Coleman and Branlard. A subsequent quantitative assessment examines discrepancies in amplitude and phase among the models, and a final yaw moment analysis highlights the practical implications of the proposed approach.

3.1 | SR-Based Yaw Engineering Model

3.1.1 | Normal Induced Velocity

Figure 5 shows the Pareto front obtained from the regression of the normal induced velocity V_{ni}. Each point on the front corresponds to a candidate SR model. The horizontal axis represents the model complexity (*Cplx*), and the vertical axis represents the loss. In this study, the loss is defined as the MSE between the model predictions and the reference ALM results for all training datasets (Tr.1, Tr.2, and Tr.3 in Table 1). The Pareto front thus illustrates the trade-off between predictive accuracy and *Cplx*, providing a basis for selecting models that achieve high predictive performance while maintaining reasonable simplicity [57].

To restrict the pool of potential models, a loss threshold is applied. Following the practice of Wang et al. [26], the threshold is set to 1.1 times the minimum loss observed in Figure 5, which occurs for the model with the highest complexity ($Cplx = 60$). The models that fall within this threshold have *Cplx* between 49 and 60. Given that the target variable Y for the SR process is the dimensionless induced velocity, normalized by the local annulus-averaged induced velocity, the resulting loss does not directly quantify the prediction error on a consistent physical scale. To verify that the models inside the threshold achieve satisfactory accuracy in physical terms, the actual RMSE is computed using the absolute induced velocity, that is, $Y_{\text{SR}} \cdot \overline{V_{ni}}$ and $Y_{\text{ALM}} \cdot \overline{V_{ni}}$, and then normalized by the total range of the normal induced velocity across all training datasets, denoted as ΔY_{ALM}. This normalized RMSE provides a physically interpretable error measure and confirms that the model selected under the 1.1 times minimum-loss threshold retain sufficient accuracy. The normalized RMSE is computed as

$$\begin{aligned}\text{RMSE}_{\text{norm}}(V_{ni}) &= \frac{1}{\Delta V_{ni,\text{ALM}}}\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\left(Y_{\text{SR}}^{(i)} - Y_{\text{ALM}}^{(i)}\right)\overline{V_{ni,\text{ALM}}}^{(i)}\right]^2},\\ \Delta V_{ni,\text{ALM}} &= \max(V_{ni,\text{ALM}}) - \min(V_{ni,\text{ALM}})\end{aligned} \tag{15}$$

where Y_{SR} and Y_{ALM} are the target variable from the SR model and ALM simulations, respectively, and N is the number of total data points. The resulting $\text{RMSE}_{\text{norm}}$ lies between 1.51% to 1.53% relative to ΔY_{ALM}, confirming that all models within the loss threshold are sufficiently accurate. Within this set of models, the one exhibiting the largest fractional drop f_d, marked by the green dot in Figure 5, is selected as the final model. The selected model, given in Equation (16), has a *Cplx* of 52. The *Cplx* reported here represents the *Cplx* perceived by the regression process, as discussed in Section 2.3, while the true *Cplx* is 75. For comparison, Coleman's model has a *Cplx* of 13, and Branlard's model has a *Cplx* of 65. This demonstrates that a reasonable balance between model accuracy and *Cplx*

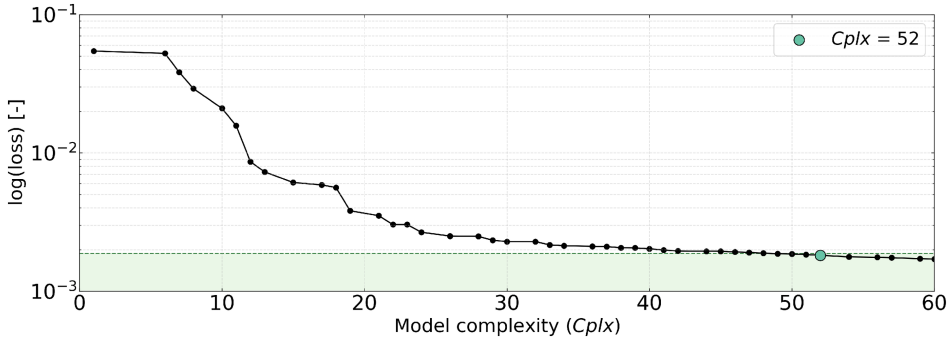

FIGURE 5 | Variation of logarithmic loss with model *Cplx* for normal induced velocity V_{ni}. The black curve denotes the Pareto front, with dots representing individual models. The region below the loss threshold is highlighted in green, and the selected model is marked prominently.

is achieved. The same convention is applied to the subsequent model of tangential induced velocity.

$$\frac{V_{ni}}{\overline{\overline{V}}_{ni}} = \sin\theta \tan(\gamma/2)\left[\frac{106.8\,\tilde{r}}{\tilde{r}+0.37+\overline{C}_t^{\,2}}(\tilde{r}-0.57)^3+\tilde{r}^3\right] - \cos\theta \tan(\gamma/2)\frac{(\tilde{r}-0.035/\tilde{r})^3+2.05(\cos\gamma-\tilde{r})^3}{\cos^2\theta+\cos\theta \sin\gamma \overline{C}_t+1.46}+1 \quad (16)$$

To evaluate the generalization performance of the selected model, it is first tested on previously unused datasets corresponding to case Ue.1 in Table 2. The $\mathrm{RMSE}_{\mathrm{norm}}$ for each yaw angle in case Ue.1 is reported in Table 3. In this context, ΔY_{ALM} in Equation (15) is replaced by the range of Y_{ALM} for each specific yaw angle rather than the range across all yaw angles. The results show that for all unseen yaw angles in case Ue.1, the model achieves an error below 3%, indicating reliable predictive performance. Moreover, for positive and negative yaw angles of the same magnitude, the $\mathrm{RMSE}_{\mathrm{norm}}$ is identical, demonstrating that the model accurately captures the symmetry between positive and negative yaw angles. This symmetry occurs because, under uniform inflow, negative yaw angles produce induced velocities identical in magnitude to positive yaw angles, differing only in phase by 180° at each annulus, as also observed by Bastankhah and Porté-Agel [60]. Further verification is therefore performed using only positive yaw angles, which are sufficient to capture the full behavior across the yaw angle range in this study.

TABLE 3 | $\mathrm{RMSE}_{\mathrm{norm}}$ of the SR model relative to reference ALM results in predicting the normal induced velocity (V_{ni}) across yaw angles (case Ue.1).

γ (°)	−25	−15	−5	5	15	25
$\mathrm{RMSE}_{\mathrm{norm}}$ (%)	2.12	1.68	0.93	0.93	1.68	2.12

Having established the model's overall predictive performance, a more detailed examination of its spatial behavior is now presented. The contour distributions of the normal induced velocity V_{ni} and the corresponding local absolute error of the SR model relative to the ALM results for yaw angles $\gamma = 5°$ and $\gamma = 25°$ are shown in Figure 6, excluding the root region ($r < 0.2R$). It should be noted that the SR predictions shown here correspond to model outputs using features obtained from ALM (also Section 3.1.2). It should be noted that the SR predictions shown here, along with those reported in Section 3.1.2, correspond to model outputs using features obtained from ALM. All quantities are normalized by the free-stream velocity. Across the rotor plane, the SR model successfully reproduces V_{ni}, reflecting its amplitude and phase variations. Near the blade tip, V_{ni} exhibits a quasisinusoidal variation, with maxima between $\theta = 90°$ and $\theta = 180°$ and minima between $\theta = 270°$ and $\theta = 360°$, consistent

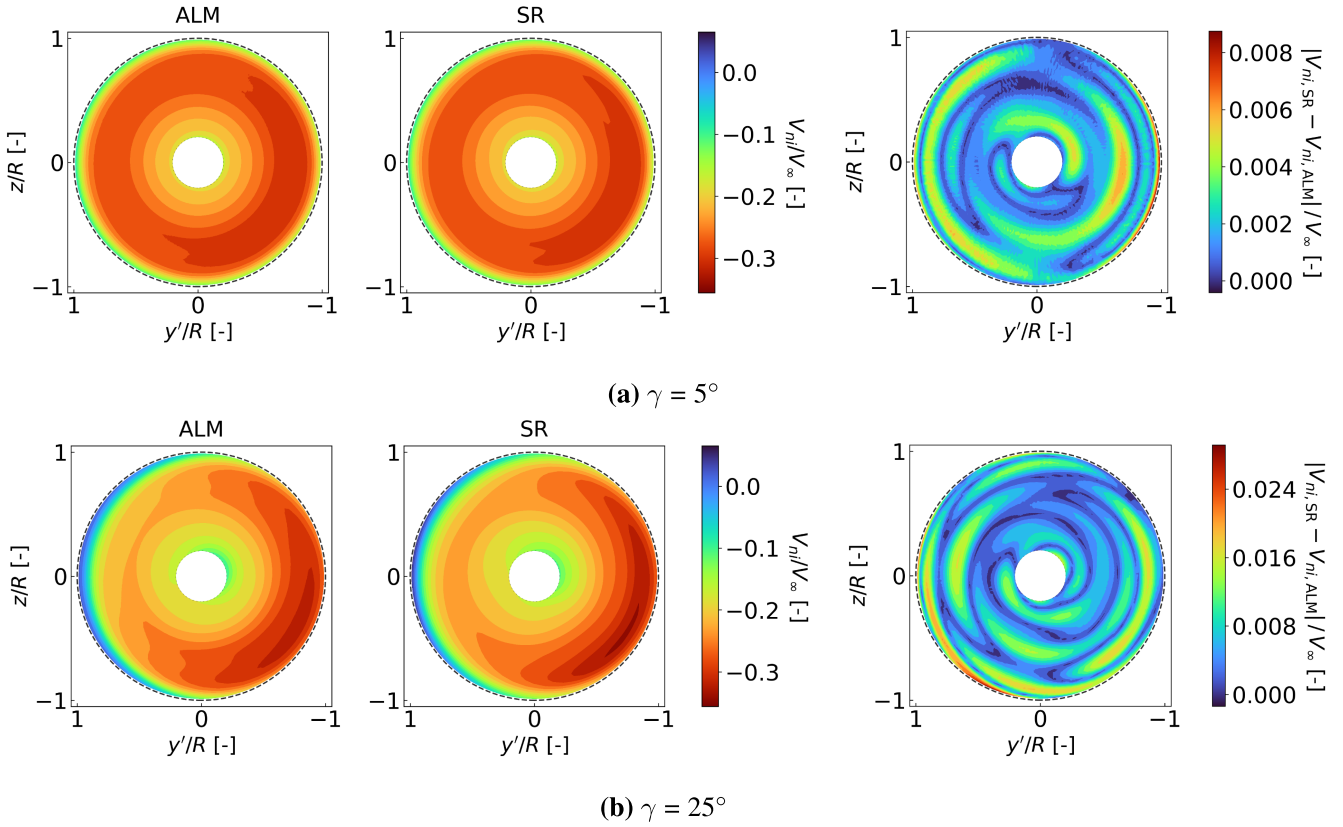

FIGURE 6 | Comparison of the normal induced velocity V_{ni} from ALM results and the selected SR model ($Cplx = 52$) at yaw angles $\gamma = 5°$ and $\gamma = 25°$. Left: Contour distributions of V_{ni}. Right: Local error magnitude distribution of the SR model relative to the ALM results. All contours are normalized by the free-stream velocity V_∞ and viewed from upstream, with the left side corresponding to the upwind side and the right side to the downwind side for positive yaw.

with previous observations [17]. Toward the midspan and blade root, the model also captures the phase shift, with the midspan maximum shifting to approximately $\theta = 180°–270°$ and the minimum to $\theta = 0°–90°$ [15]. The absolute error at each actuator point, shown in the right panel of Figure 6, indicates that deviations are small compared with the overall variation of V_{ni}. For yaw angles of 5° and 25°, the maximum deviation remains below 0.1% and 3%, respectively, while V_{ni} varies by up to 35% of the free-stream velocity. For clarity, only the results for $\gamma = 5°$ and $\gamma = 25°$ are shown. Other yaw angles exhibit similar trends and are therefore omitted. The full dataset is available in Sun et al. [61]. Overall, these results demonstrate that the SR model accurately captures the normal induced velocity across the blade span and that the model achieves reliable predictive performance while remaining compact and relatively simple.

3.1.2 | Tangential Induced Velocity

Following the same procedure as for the normal induced velocity V_{ni}, the Pareto front for the tangential induced velocity V_{ti} is shown in Figure 7. Applying the same loss threshold as in Section 3.1.1 yields a set of candidate models with *Cplx* ranging from 42 to 60. The normalized error $\mathrm{RMSE}_{\mathrm{norm}}$ is computed using Equation (17), with the tangential component substituted for the normal component. The resulting $\mathrm{RMSE}_{\mathrm{norm}}$ values lie between 3.36% and 3.54%, indicating that the SR models for V_{ti} maintain an acceptably low error level across this range of *Cplx*. The final model is selected based on the optimal f_d within the loss threshold and is given in Equation (18). Its reported *Cplx* is 48, representing the *Cplx* defined internally by the SR algorithm. However, this reported value does not account for the additional *Cplx* introduced by the preprocessed input features, such as trigonometric or normalized terms supplied to the algorithm. Accounting for these contributions, the true model *Cplx* rises to 66. For reference, Coleman's analytical model has a *Cplx* of 13, and Branlard's model has a *Cplx* of 68. These results indicate that the selected SR model maintains a low prediction error on the training datasets while keeping a *Cplx* comparable with established yaw engineering models.

$$\mathrm{RMSE}_{\mathrm{norm}}(V_{ti}) = \frac{1}{\Delta V_{ti,\mathrm{ALM}}}\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\left(Y_{\mathrm{SR}}^{(i)} - Y_{\mathrm{ALM}}^{(i)}\right)\overline{V_{ti}}^{(i)}\right]^2}, \quad \Delta V_{ti,\mathrm{ALM}} = \max(V_{ti,\mathrm{ALM}}) - \min(V_{ti,\mathrm{ALM}}) \tag{17}$$

$$\frac{V_{ti}}{\overline{V_{ti}}} = \frac{\sin\gamma}{1.74}\left[2.67(\cos\theta + \tilde{r}\sin\theta) + \frac{0.26 - 1.88\tilde{r} + (\overline{C_t} + 0.24)\,\tilde{r}^3/0.10 - (0.23 - 2\tilde{r})\,(\tan(\gamma/2) + \sin\gamma)\,\sin\theta}{\overline{C_t} - 0.05\cos\gamma}\cos\theta\right] + 1 \tag{18}$$

Similarly, the generalization performance is evaluated for the tangential induced velocity V_{ti} in case Ue.1, and the computed $\mathrm{RMSE}_{\mathrm{norm}}$ values are presented in Table 4. Across all yaw angles, the $\mathrm{RMSE}_{\mathrm{norm}}$ values remain below 5%, indicating that the model achieves satisfactory accuracy for V_{ti}. Consistent with the trend observed for V_{ni} in Table 3, the $\mathrm{RMSE}_{\mathrm{norm}}$ values for V_{ti} coincide exactly for positive and negative yaw angles of equal magnitude. This agreement reflects the symmetry previously discussed for V_{ni}, where the induced velocity for opposite yaw directions differs only in phase and not in magnitude under uniform inflow conditions. Consequently, only the positive yaw angles are considered in the subsequent verifications.

Figure 8 presents the tangential induced velocity V_{ti} and its associated error magnitude at each actuator point. The SR model

TABLE 4 | $\mathrm{RMSE}_{\mathrm{norm}}$ of the SR model relative to reference ALM results in predicting the tangential induced velocity (V_{ti}) across yaw angles (case Ue.1).

γ (°)	−25	−15	−5	5	15	25
$\mathrm{RMSE}_{\mathrm{norm}}$ (%)	4.87	4.53	3.64	3.64	4.53	4.87

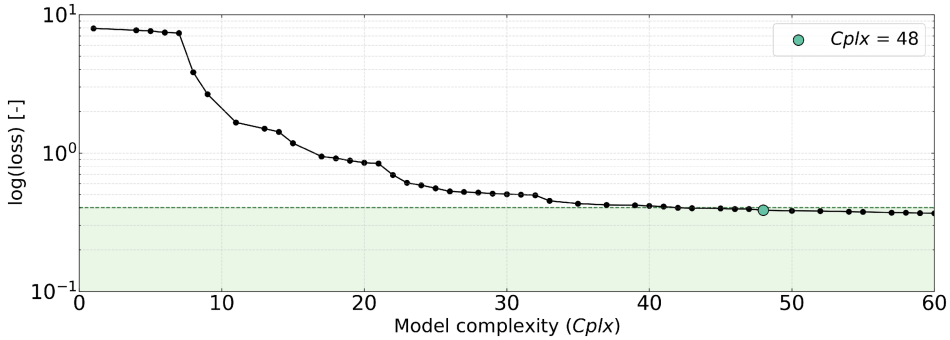

FIGURE 7 | Variation of logarithmic loss with model *Cplx* for tangential induced velocity V_{ti}. The black curve denotes the Pareto front, with dots representing individual models. The region below the loss threshold is highlighted in green, and the selected model is marked prominently.

reproduces the azimuthal variation of V_{ti} with high accuracy. In contrast to V_{ni}, which exhibits a phase shift from blade tip to root, V_{ti} maintains a consistent quasicosine variation along the span, which is well captured by the model. The local error magnitude remains below 0.3% for $\gamma = 5°$ and 1.5% for $\gamma = 25°$, while the variation of V_{ti} itself reaches up to around 12%, indicating that the deviation is acceptably small. These results demonstrate that the selected SR model reliably predicts the tangential induced velocity across the rotor plane.

3.2 | Model Verification in BEM Simulation

3.2.1 | Induction and Forces

This section evaluates the performance of selected SR models in Section 3.1 within the BEM framework. Yaw engineering models proposed by Coleman, Branlard, and the novel SR model developed in this study are implemented in BEM, and their predictions are compared with ALM numerical

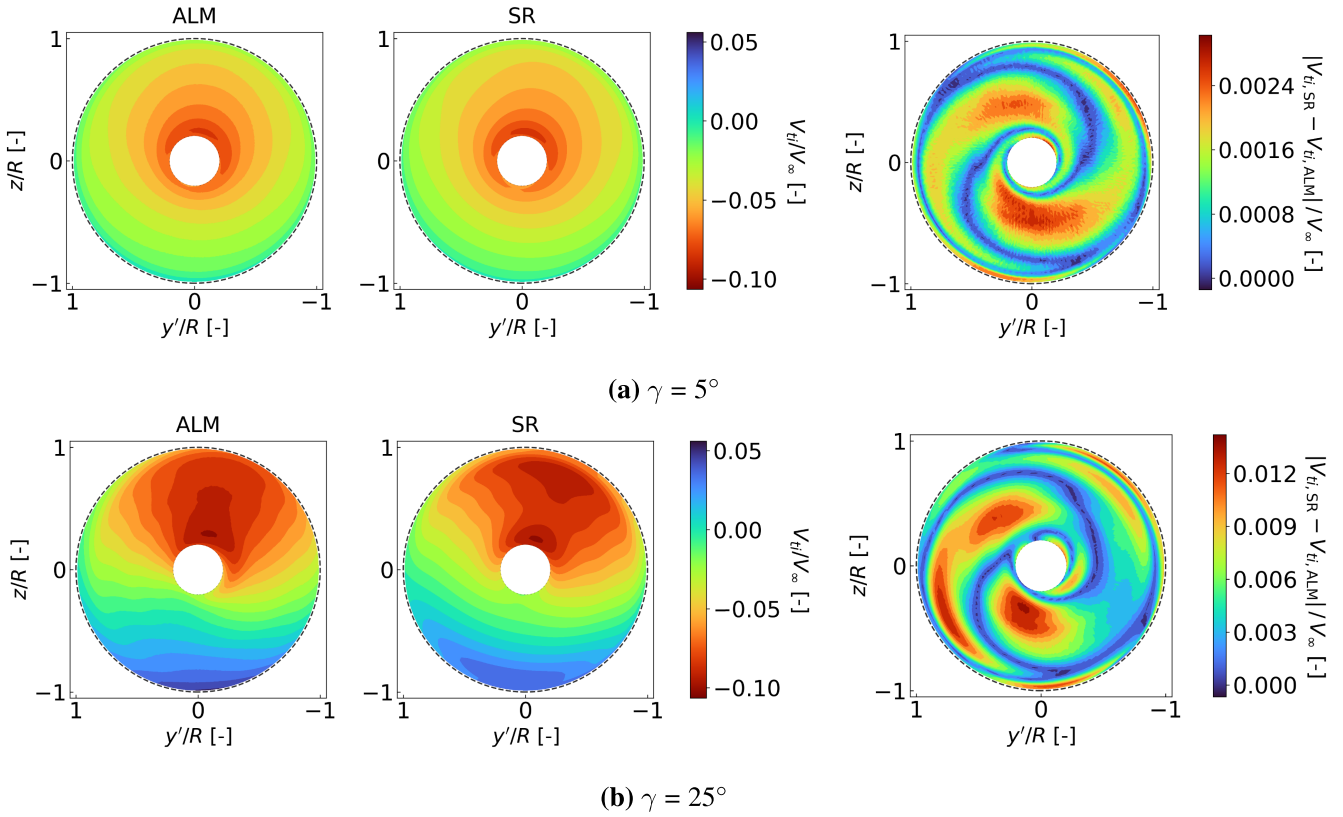

FIGURE 8 | Comparison of the tangential induced velocity V_{ti} from ALM results and the selected SR model ($Cplx = 48$) at yaw angles $\gamma = 5°$ and $\gamma = 25°$. Left: Contour distributions of V_{ti}. Right: Local error magnitude distribution of the SR model relative to the ALM results. All contours are normalized by the free-stream velocity V_∞ and viewed from upstream, with the left side corresponding to the upwind side and the right side to the downwind side for positive yaw.

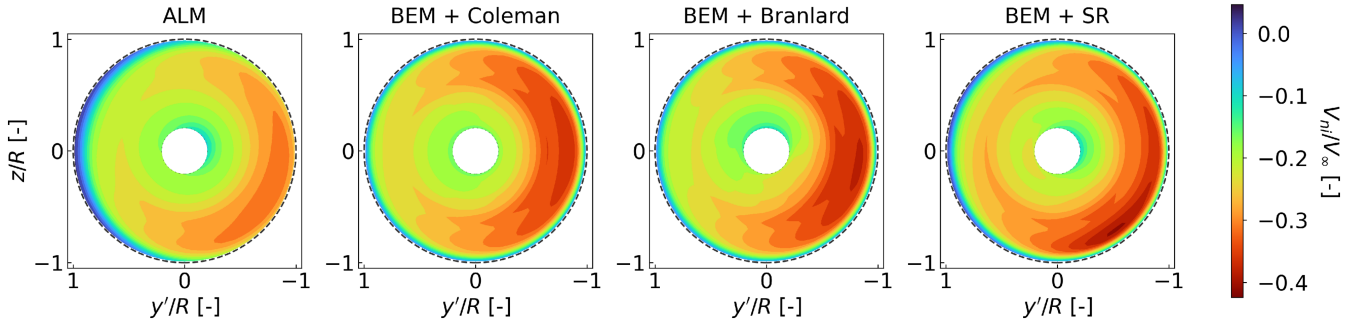

FIGURE 9 | Contour plot of the normal induced velocity V_{ni} across the rotor plane for the ALM and BEM simulations using the Coleman, Branlard, and SR models at a yaw angle of $\gamma = 25°$ (viewed from upstream).

simulations for case Ue.1 at a yaw angle of $\gamma = 25°$. The analysis focuses on the azimuthal variations of induced velocities and the resulting aerodynamic forces along the main blade span. Before proceeding with the detailed analysis, it is worth noting that the overall amplitudes of both induced velocities and forces predicted by the BEM models exhibit a slight offset compared with the ALM simulations. This discrepancy arises from inherent differences between the BEM and ALM methodologies and lies beyond the scope of the present study.

Figure 9 shows the disk contour of the normal induced velocity distribution, excluding the root region ($r < 0.2R$). All models capture the higher induction on the downwind side near the blade tip. Moving inward along the blade span, differences among the three models become more pronounced. Coleman's model maintains the higher induction on the downwind side across the span, but its deviation from the ALM results increases toward the blade root, as it fails to capture the phase shift along the annulus observed in the ALM data. Branlard's model introduces a modest phase shift below midspan, whereas the SR model closely follows the ALM phase, showing only a minor amplitude overshoot near the blade root within the azimuthal sector 180°–270°.

Figures 10 and 11 provide a more detailed view of the azimuthal distributions of normal induced velocity V_{ni} and normal force f_n at eight radial locations along the blade span. In the midspan to tip region ($r = 0.6R$–$0.9R$), while all models show similar trends near the blade tip, the SR model shows a slight but consistent improvement in phase agreement

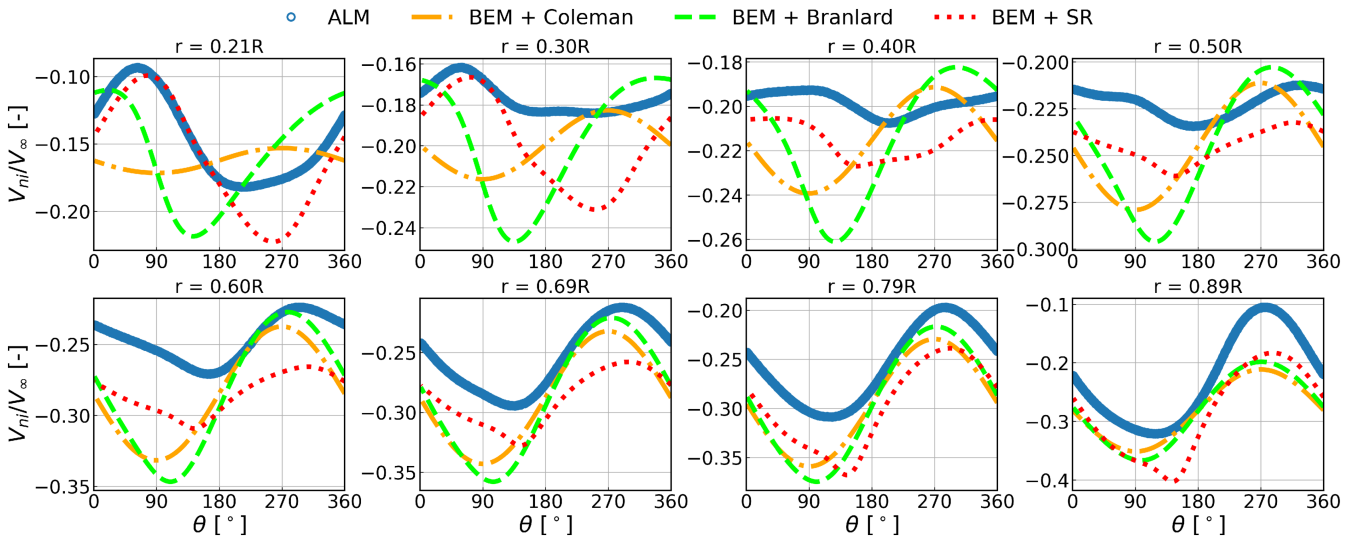

FIGURE 10 | Azimuthal distributions of normal induced velocity V_{ni} at eight radial positions ($r = 0.2R$–$0.9R$) for $\gamma = 25°$. Each subplot represents a radial location, with ALM results as reference and overlaid predictions from BEM simulations using the Coleman, Branlard, and SR models for comparison.

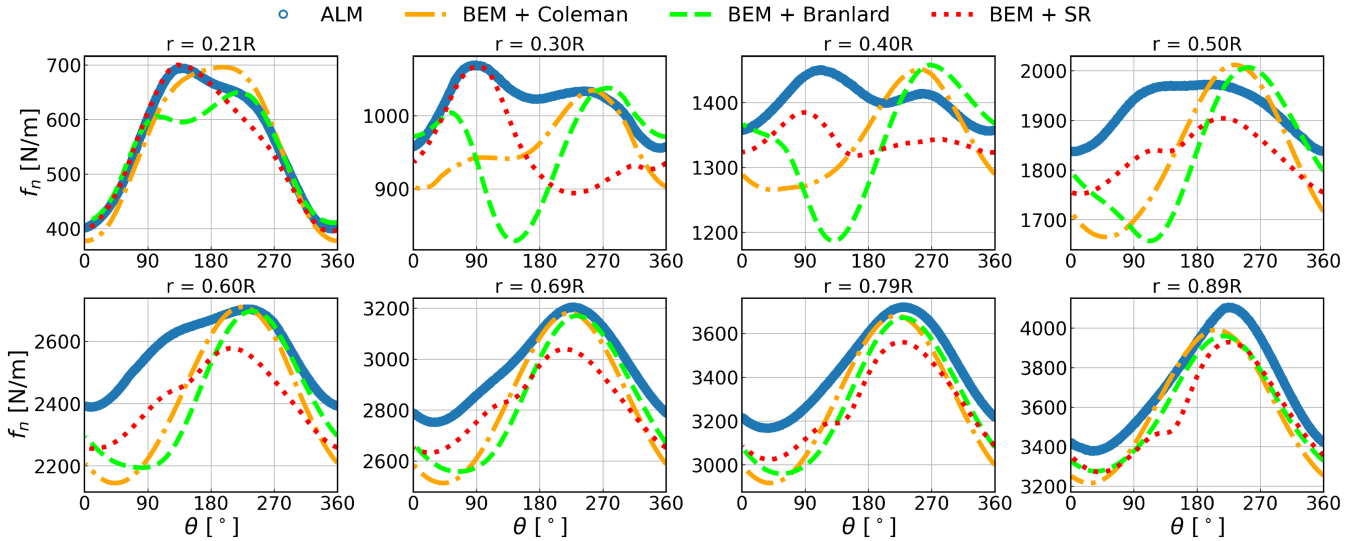

FIGURE 11 | Azimuthal distributions of normal forces F_n at eight radial positions ($r = 0.2R$–$0.9R$) for $\gamma = 25°$. Each subplot represents a radial location, with ALM results as reference and overlaid predictions from BEM simulations using the Coleman, Branlard, and SR models for comparison.

compared with the other two models when referenced to the ALM results. This improved phase alignment is reflected in both the induced velocity and aerodynamic force. The region from midspan to root ($r = 0.2R$–$0.5R$) is of particular interest, as it exhibits the most pronounced discrepancies among the three models. Owing to the increasing influence of the root vortex [17], a clear phase shift in the induced velocity distribution is observed, distinguishing this region from the tip region. Consequently, Coleman's model becomes unreliable here, as it considers only the sinusoidal variation induced by the tip vortex and neglects the effect of the root vortex. Branlard's model improves the phase alignment of induced velocity in this region by incorporating the root vortex as shown in Figure 10. Nevertheless, a residual phase error persists throughout this region. In addition, a significant amplitude offset occurs between the 90° and 180° azimuthal angles. This amplitude discrepancy is attributed to an overestimation of the root vortex contribution under the assumption of uniform disk circulation introduced in Section 1, which is consistent with the observations reported in Branlard's work [15]. In the corresponding force distribution shown in Figure 11, ALM data show a maximum force near the 90° azimuthal position at radial locations between $r = 0.2R$ and $r = 0.4R$, whereas both Coleman's and Branlard's models predict the peak around 250°. This phase misalignment of forces directly results from the misprediction of induced velocity shown in Figure 10, where both Coleman's and Branlard's models overestimate the induced velocity around $\theta = 250°$ and underestimate it near $\theta = 90°$. However, the SR model effectively resolves this misalignment. As demonstrated in Figures 10 and 11, it achieves consistent phase alignment across all radial positions, correcting the discrepancies observed in earlier models, especially between $r = 0.2R$ and $r = 0.4R$. Although a slight amplitude underprediction occurs at $r = 0.3R$, the SR model demonstrates superior overall performance among the three yaw engineering models, showing both consistent phase alignment and amplitude predictions that match the ALM data. Therefore, although all three models perform well at the outboard section, the SR model exhibits slightly better phase alignment there and demonstrates superior performance toward the inboard section, highlighting its overall reliability in predicting the normal induced velocity under yawed conditions.

Figures 12–14 compare the tangential induced velocity V_{ti} and tangential force f_t predicted by the three BEM models with the ALM results. Figure 12 shows the contour distribution of V_{ti}, providing an overall view of the induced velocity field, while Figures 13 and 14 present the azimuthal distributions of tangential induced velocity and tangential force at the same radial positions analyzed for the normal induced velocity. Unlike the normal induced velocity, the tangential induced velocity exhibits a quasicosine variation at each annulus with no apparent phase shift along the blade span. The contour in Figure 12 indicates that Coleman's model deviates significantly from the ALM results, making it unreliable and showing that the assumption in the literature [1] that the tangential induced velocity follows the same sinusoidal variation as the normal induced velocity is not valid. In contrast, both Branlard's model and the SR model reproduce the overall induced velocity distribution observed in the ALM results. The azimuthal distributions in Figure 13 show that Branlard's model produces amplitudes slightly higher than those of the ALM along the blade span, with a noticeable drop near the root to midspan at $r = 0.2R$–0.4R around $\theta = 90°$, likely caused by the overestimation of the root vortex, as also observed for the normal induced velocity in Figure 10. The SR model, in comparison, reproduces the amplitudes more consistently with the ALM data, showing smaller deviations throughout the span. A trend similar to that observed for the normal force is evident in the tangential force distribution shown in Figure 14, where all three models show good agreement in amplitude and phase from midspan to tip. As the radial position approaches the root, phase deviations become more apparent in both Coleman's and Branlard's models, which aligns with the discrepancies identified in the normal force results. It should be noted that the aerodynamic forces result from the projection of lift (L) and drag (D) onto the normal and tangential directions, and the variation in normal induced velocity plays a dominant role in shaping the velocity triangle in Figure 4, which governs the distribution of the forces. As a result, although the tangential induced velocity differs noticeably among the models, particularly for Coleman's model, the tangential force profile remains relatively consistent. Consequently, a similar phase shift in the location of the maximum is observed as in the normal force distribution in Figure 11. Therefore, we conclude that the SR model for tangential induced velocity demonstrates the most robust performance among the three BEM models.

Because the model was both regressed (cases Tr.1, Tr.2, and Tr.3) and verified (case Ue.1) using data under the NREL 5-MW

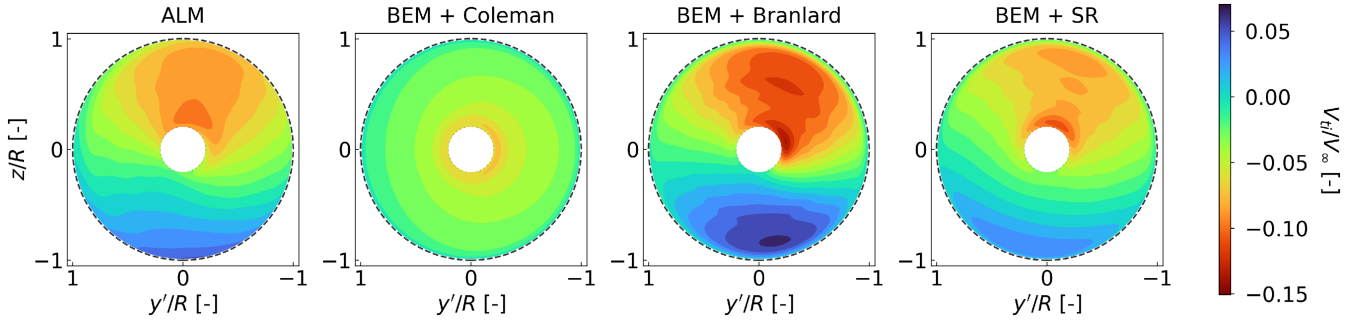

FIGURE 12 | Contour plot of the tangential induced velocity V_{ti} across the rotor plane for the ALM and BEM simulations using the Coleman, Branlard, and SR models at a yaw angle of $\gamma = 25°$ (viewed from upstream).

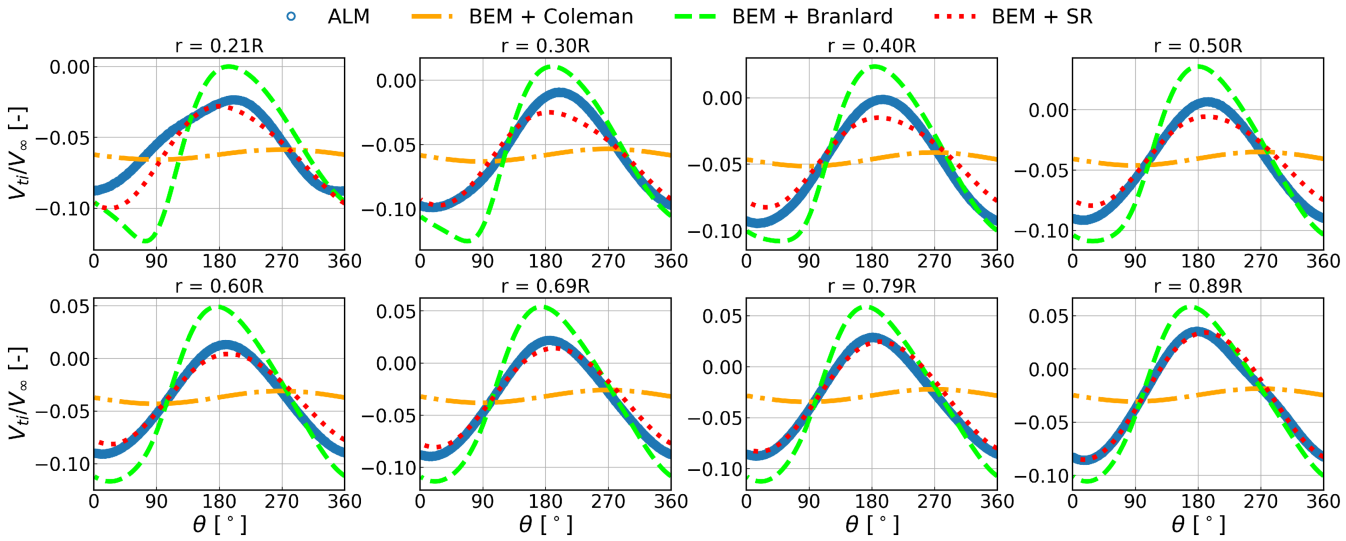

FIGURE 13 | Azimuthal distributions of tangential induced velocity V_{ti} at eight radial positions ($r = 0.2R$–$0.9R$) for $\gamma = 25°$. Each subplot represents a radial location, with ALM results as reference and overlaid predictions from BEM simulations using the Coleman, Branlard, and SR models for comparison.

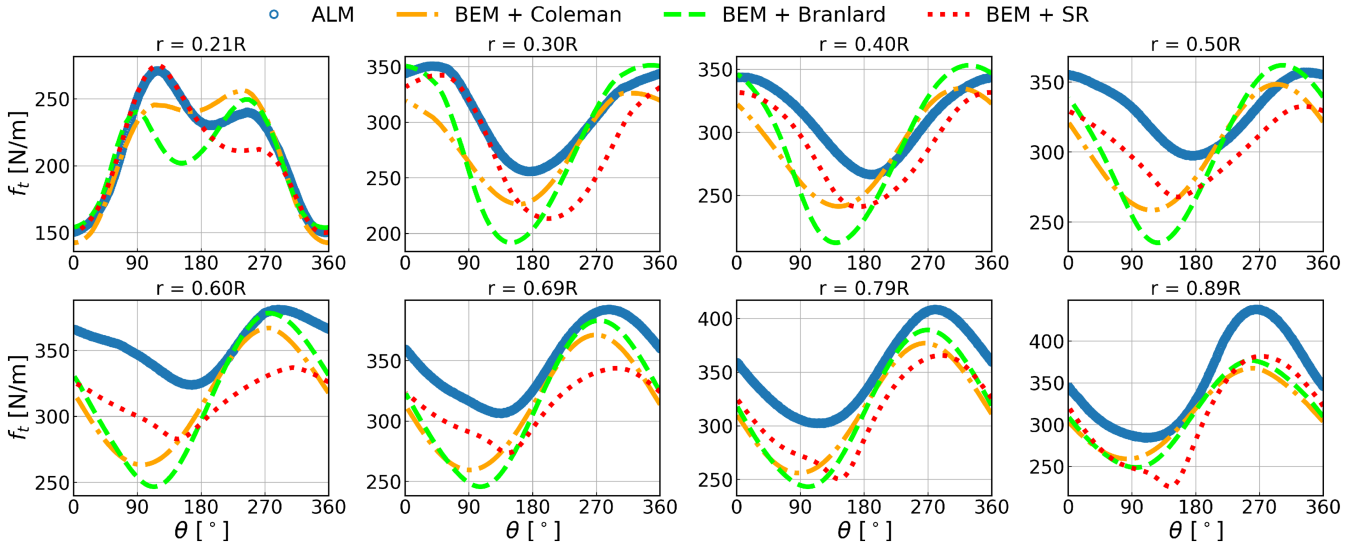

FIGURE 14 | Azimuthal distributions of tangential forces f_t at eight radial positions ($r = 0.2R$–$0.9R$) for $\gamma = 25°$. Each subplot represents a radial location, with ALM results as reference and overlaid predictions from BEM simulations using the Coleman, Branlard, and SR models for comparison.

below-rated operating conditions ($V_\infty = 8.0$ m/s, $\lambda = 7.55$), its robustness was further evaluated through two additional cases: the same wind turbine at rated wind speed (case Ue.2, $V_\infty = 11.4$ m/s, $\lambda = 7.0$) and an independent IEA 15-MW wind turbine (case Ue.3, $V_\infty = 9.0$ m/s, $\lambda = 9.0$), as summarized in Table 2. The findings from these additional tests corroborate the previous conclusions and are presented in Appendices C and D. Overall, the results indicate that the SR model developed in this study reliably captures the phase, maintains reasonable amplitude across the blade span, and consistently outperforms both Coleman's and Branlard's models.

3.2.2 | Quantitative Error Analysis

To further verify the model, a quantitative error analysis based on the amplitude and phase differences between the ALM results and the BEM model predictions is conducted in this section. Figure 15 presents the RMSE of the normal and tangential induced velocities between the BEM predictions of the three models and the ALM results, evaluated over yaw angles from − 30° to 30° for the NREL 5-MW turbine in case Ue.1, which is used for model training and testing. All RMSE values are normalized by the free-stream velocity in this analysis. Figure 16

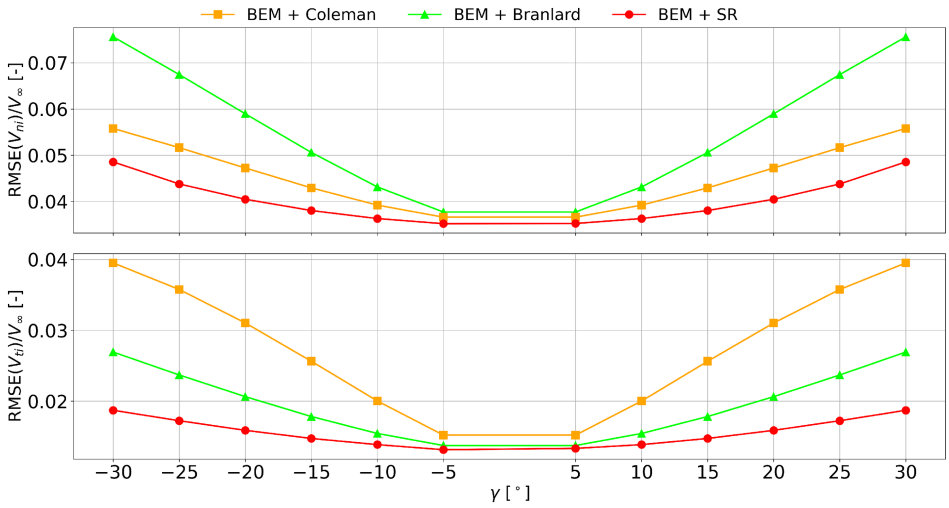

FIGURE 15 | Comparison of V_{ni} and V_{ti} RMSE between three yaw engineering models and ALM results across all yaw angles in case Ue.1. Top: Normal induced velocity. Bottom: Tangential induced velocity.

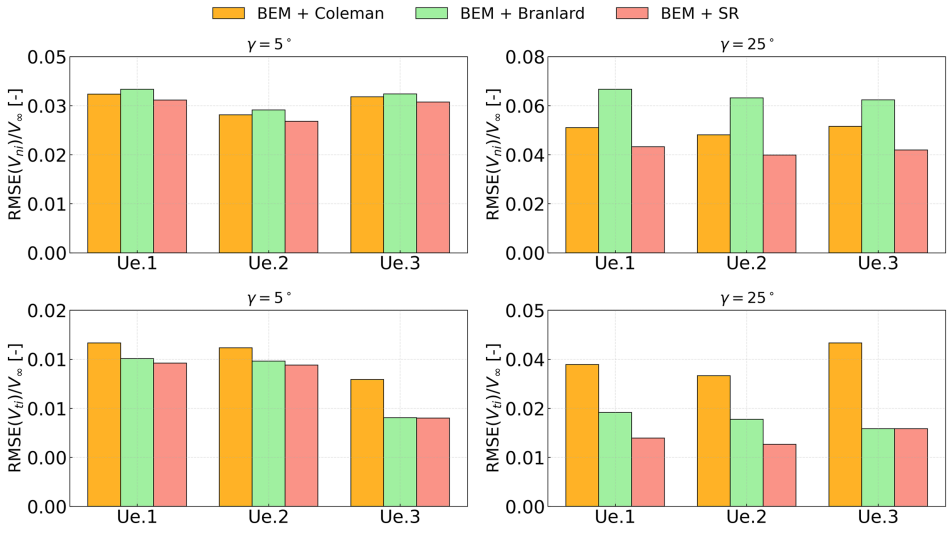

FIGURE 16 | Comparison of V_{ni} and V_{ti} RMSE between three yaw engineering models and ALM results across three wind turbine operating cases for $\gamma = 5°$ and 25°. Cases include Ue.1, Ue.2, and Ue.3.

summarizes the corresponding normalized RMSE values for the three turbine operating conditions at yaw angles of 5° and 25° in cases Ue.1, Ue.2, and Ue.3. Although a consistent discrepancy exists between BEM and ALM, this discrepancy influences all engineering models in a similar manner. Despite this discrepancy, the SR model consistently yields the lowest errors across all operating conditions, demonstrating its robustness and general applicability.

Figure 17 further examines the phase differences of the induced velocity at each radial annulus, where the x axis denotes the normalized radial position and the y axis shows the phase difference between BEM and ALM results. These phase differences were computed based on cross-correlation analysis to accurately quantify the relative shifts between the two simulations, as introduced in Section 2.4. The shaded band represents the variation across all yaw angles in each case, while the dashed line indicates the mean value. Both the training and unseen turbine cases (Tr.1 and Ue.1) across all yaw angles and the two additional unseen cases (Ue.2 and Ue.3) at 5° and 25° yaw are considered, consistent with the RMSE analysis. For the normal induced velocity, all three models show good phase agreement near the blade tip, consistent with the observations in Section 3.2.1, with the SR model exhibiting the least phase difference. Toward the blade root, phase differences become more pronounced. Coleman's model exhibits the largest misalignment, as it relies solely on a single sinusoidal variation, thereby resulting in a maximum phase shift of approximately 175°. Branlard's model improves the prediction by accounting for the complete vortex system, reducing the maximum deviation to about 125°. The SR model performs best, maintaining relatively small phase differences, with a maximum shift of only around 20°–40° between $0.2R$ and $0.6R$. For the tangential induced velocity, Coleman's model exhibits substantial discrepancies because it is not applicable in this scenario as discussed in section 3.2.1, whereas Branlard's

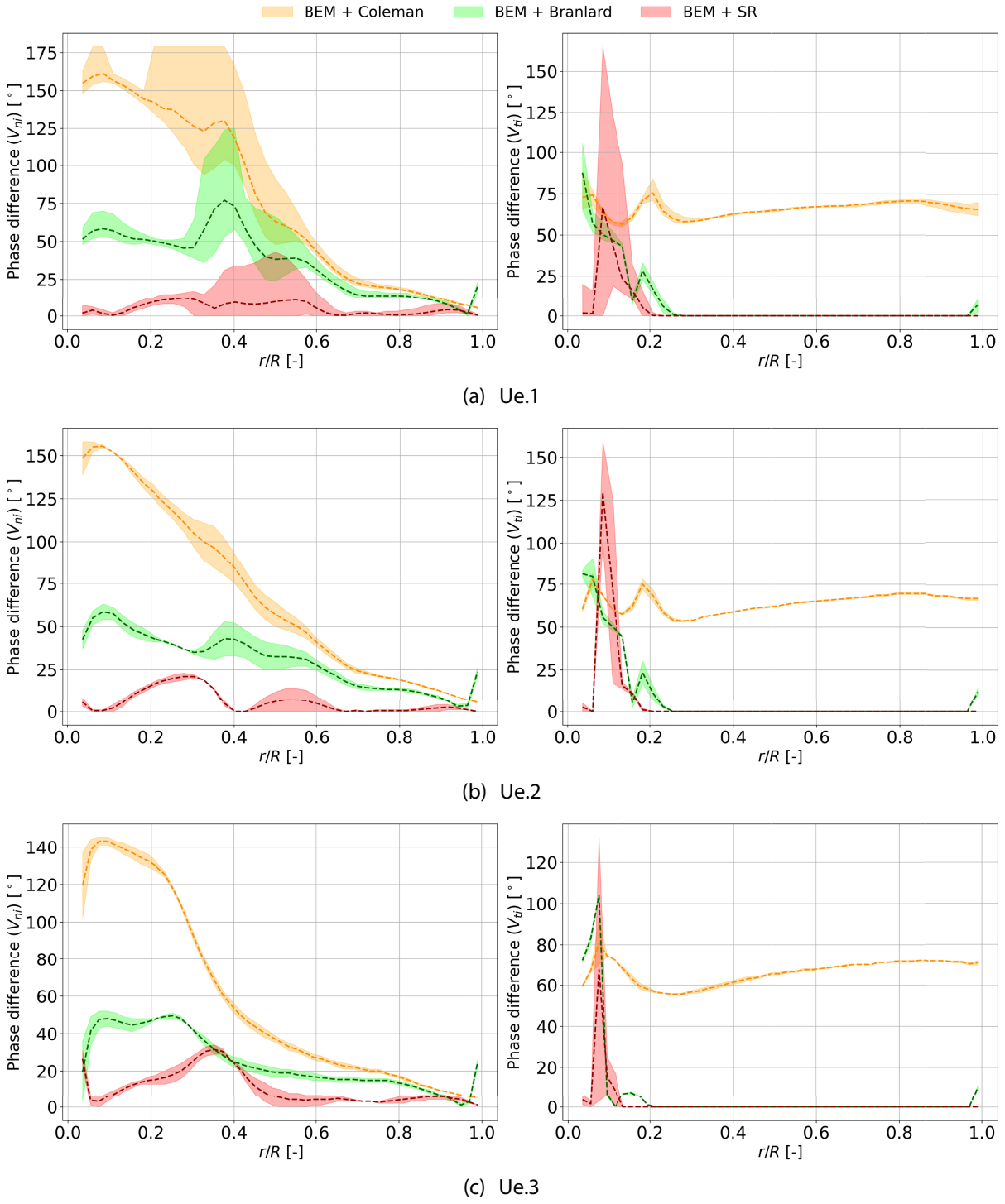

FIGURE 17 | Comparison of phase differences of normal and tangential induced velocities at each radial annulus between three yaw engineering models and ALM results across all yaw angles along the blade span. Left: Normal induced velocity. Right: Tangential induced velocity. Shaded regions show the full yaw angle variation, with dashed lines denoting mean values.

model and the SR model accurately capture the phase over most of the span and remain well aligned overall, except near the blade root ($r < 0.2R$). These results, supported by both qualitative and quantitative assessments across a range of operational conditions and turbine types, demonstrate that the SR model proposed in this study provides a robust capability for accurate, low-fidelity prediction of induced velocities in wind turbines operating under yawed conditions.

3.2.3 | Yaw Moment Analysis

As discussed in Section 3.2.1, wind turbine yaw misalignment introduces a phase shift in the induced velocities and resulting aerodynamic forces across the rotor plane, thereby modifying the load distribution and potentially exerting a significant influence on the resulting yaw moment [59]. In this context, this section further examines the impact of different yaw engineering models

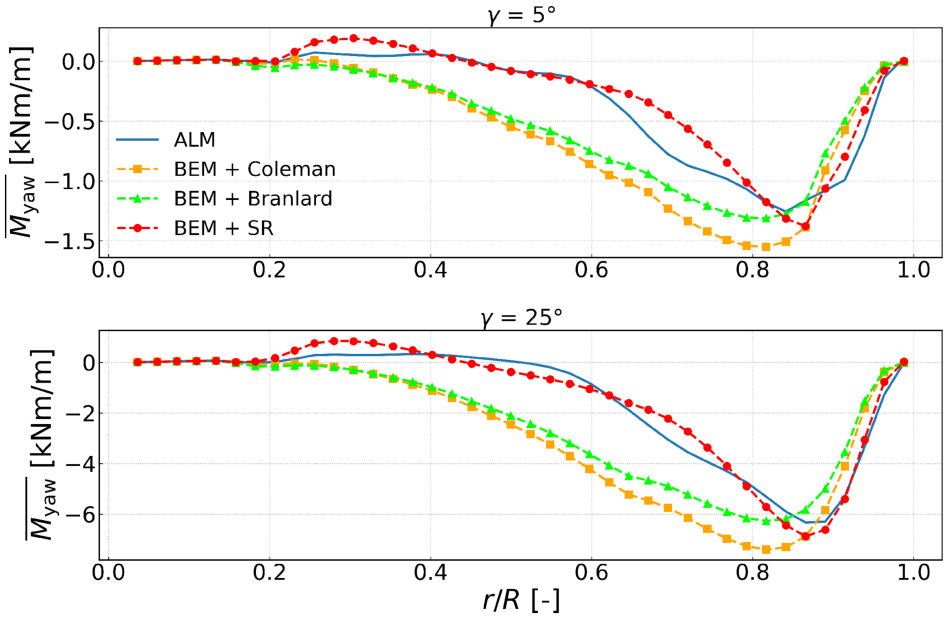

FIGURE 18 | Comparison of annulus-averaged $\overline{M_{yaw}}$ between three yaw engineering models and ALM results along the blade span. Top and bottom panels correspond to yaw angles of $\gamma = 5°$ and 25°, respectively (case Ue.1).

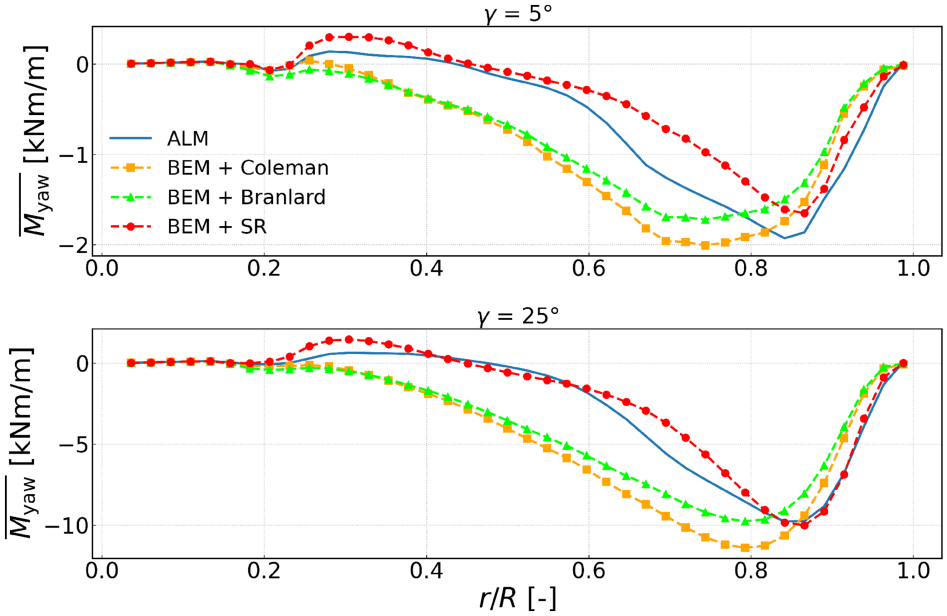

FIGURE 19 | Comparison of annulus-averaged $\overline{M_{yaw}}$ between three yaw engineering models and ALM results along the blade span. Top and bottom panels correspond to yaw angles of $\gamma = 5°$ and 25°, respectively (case Ue.2).

on loading prediction by examining the yaw moment ($\overline{M_{yaw}}$), averaged azimuthally at each radial annulus. Figures 18–20 compare the radial distribution of $\overline{M_{yaw}}$ from ALM simulations with BEM predictions from the three yaw engineering models for yaw angles of $\gamma = 5°$ and $\gamma = 25°$ across the three unseen cases. In the ALM simulations, the rotor generates a beneficial restoring yaw moment above the midspan region ($r/R > 0.5$) that tends to realign the turbine with the inflow, whereas below midspan, the yaw moment reverses due to the phase shift of peak aerodynamic forces, resulting in a nonrestoring yaw moment at the inboard section of the rotor [1], as discussed in Section 3.2.1. Coleman's and Branlard's models qualitatively capture the restoring yaw moment but show increasing deviation from the ALM results toward the inboard section of the rotor. Owing to the misalignment in the predicted force phase near midspan, both models fail to reproduce the reversal in yaw moment direction and instead predict a restoring moment across the entire rotor plane. In contrast, the SR model exhibits markedly better agreement with the ALM results, capturing both the restoring yaw moment above midspan and the opposing nonrestoring yaw moment below midspan. Furthermore, it shows improved correspondence with the ALM results in terms of the yaw moment

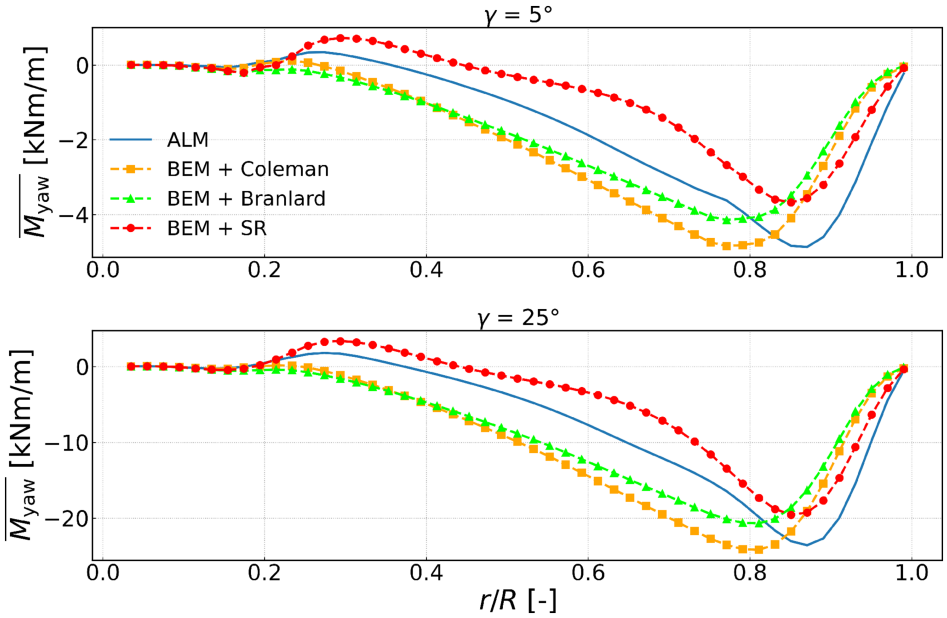

FIGURE 20 | Comparison of annulus-averaged $\overline{M_{yaw}}$ between three yaw engineering models and ALM results along the blade span. Top and bottom panels correspond to yaw angles of $\gamma = 5°$ and 25°, respectively (case Ue.3).

amplitude. Therefore, the SR model provides a more accurate representation of the radial yaw moment distribution and is better suited for predicting yaw-induced loading on the rotor.

4 | Conclusions

This study develops a pair of novel yaw engineering models, derived from ALM simulations using SR, to capture the nonuniformity in the normal and tangential induced velocities resulting from the skewed wake effect of a static yawed wind turbine. The models are trained on ALM simulation data of the NREL 5-MW turbine, encompassing a range of yaw angles and three distinct pitch settings, thereby covering a wide span of thrust coefficients (C_T) within which the skewed wake effect plays a dominant role in determining the local angle of attack and the resulting aerodynamic loading. The models are comparatively assessed via comparison with existing models (Coleman's and Branlard's), using the ALM result as reference. The proposed model offers two key advantages. First, the induced velocities and loading predicted by the SR model exhibit closer agreement with the reference ALM results than Coleman's and Branlard's models, with the most notable improvements being its ability to capture the phase shift of the aerodynamic force distribution from blade tip to root. Second, the newly developed model strikes a balance between accuracy and model complexity, maintaining a reasonable level of complexity that is compact and user-friendly for wind turbine simulation applications. Several conclusions can be drawn from the results presented.

The SR model consistently outperforms existing yaw engineering models across the blade span. For the normal induced velocity, a slight improvement in phase alignment is observed with the SR model at the blade tip. Significant discrepancies occur between the midspan and blade root due to the increasing influence of the root vortex. Coleman's and Branlard's models show notable deviations in predicting the location of maximum forces, whereas the SR model accurately captures both phase and spatial distribution along the blade. For the tangential induced velocity, Coleman's model proves unreliable, while Branlard's and the SR models show good agreement, with the SR model providing a slightly closer match in amplitude. Similar phase misalignments are observed in the tangential force distributions. Quantitative error analysis further demonstrates that the SR model consistently exhibits the lowest prediction errors and smallest phase discrepancies among the three models for both velocity components.

The yaw moment analysis demonstrates that the SR model clearly outperforms both Coleman's and Branlard's models in predicting the yaw moment distribution along the rotor radius. Unlike the Coleman's and Branlard's models, it successfully captures the reversal of the yaw moment below the midspan, which results from the spanwise phase shift of peak aerodynamic forces, and exhibits a closer match in amplitude. This improvement highlights the SR model's ability to reproduce the key physical features of yaw-induced loading and indicates that it provides a more reliable and physically consistent representation of rotor aerodynamics under yawed inflow conditions.

It is important to emphasize that a careful preprocessing for SR, including feature selection and operator constraints, is essential to ensure the discovery of expressions. In this study, all features and the target variable are expressed in dimensionless form to preserve physical consistency. Features such as the azimuthal angle (θ) and yaw angle (γ) are transformed using trigonometric functions as input data to avoid unphysical terms such as $\sin(r/R)$, and the use of trigonometric operators is prohibited during the regression process. Model selection is guided by the Pareto front, balancing accuracy and complexity, while a predefined threshold is applied to constrain the candidate models to those achieving a sufficient level of accuracy.

The new yaw engineering models proposed in this study have been regressed over a wide range of yaw angles and thrust coefficients, covering typical operating conditions within which the skewed wake effect strongly influence rotor aerodynamics. Extensive verification across this broad range further demonstrates the robustness of the model and its ability to accurately predict aerodynamic loads for turbines operating under a variety of conditions within these limits. These results indicate that the new set of models provide a reliable framework for capturing key aerodynamic behaviors across the full span of commonly encountered operating scenarios.

Nomenclature

A	rotor swept area (m2)
D	rotor diameter (m)
R	rotor radius (m)
r	radial position (m)
$\tilde{r}$	normalized radial position (-)
V_∞	free-stream velocity (m/s)
θ	azimuthal angle (°)
γ	yaw angle (°)
ϕ	inflow angle (°)
β	twist angle (°)
α	angle of attack (°)
V_{rel}	relative velocity (m/s)
Ω	rotational speed (rpm)
λ	tip-speed ratio (-)
V_i	induced velocity (m/s)
V_n	normal velocity (m/s)
V_{ni}	normal induced velocity (m/s)
$\overline{V_{ni}}$	annulus-averaged normal induced velocity (m/s)
V_t	tangential velocity (m/s)
V_{ti}	tangential induced velocity (m/s)
$\overline{V_{ti}}$	annulus-averaged tangential induced velocity (m/s)
F_r	Glauert model's radial dependency (-)
$A_{1,i}, A_{2,i}$	Fourier amplitude (-)
$\psi_1,\ \psi_2$	Fourier phase (rad)
C_T	rotor-averaged thrust coefficient (-)
C_t	local thrust coefficient (-)
$\overline{C_t}$	annulus-averaged thrust coefficient (-)
f_n	normal force (N/m)
f_t	tangential force (N/m)
$\overline{M_{\text{yaw}}}$	annulus-averaged yaw moment (kNm/m)
f_d	fraction drop (-)
$Cplx$	model complexity (-)
X	symbolic regression feature (-)
Y	symbolic regression target variable (-)
x, y, z	computational domain coordinates (-)
$x',\ y',\ z'$	rotor plane coordinates (-)

Author Contributions

H.S. developed the methodology, carried out the simulations, performed data analysis, and wrote the paper. A.S. provided scientific supervision throughout the analysis phase and revised the manuscript. W.Y. contributed toward funding acquisition, sharing her expertise on vortex wake modeling, scientific supervision throughout the analysis phase, and revising the manuscript.

Acknowledgments

This research is supported by the DIAMOND (DynamIc yAw Models fOr wiND turbine/farm design) project funded by the Dutch Research Council (NWO), Netherlands, through Open Technology Programme (OTP) under Grant Agreement No. 20052. The authors thank the Delft High-Performance Computing Centre [62] for the use of computational resources of the DelftBlue supercomputer. The authors thank SURF (www.surf.nl) for the support in using the Dutch National Supercomputer Snellius. The authors would also like to thank YuanTso Li, Nirav Dangi, Shantanu Purohit, and Renzhi Tian for their helpful discussions.

Funding

This research is supported by the DIAMOND (DynamIc yAw Models fOr wiND turbine/farm design) project funded by the Dutch Research Council (NWO), Netherlands, through Open Technology Programme (OTP) under Grant Agreement No. 20052.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are openly available in 4TU at https://doi.org/10.4121/1258f8d9-a8a7-4912-9187-6bc35b217e19.

References

1. M. Hansen, *Aerodynamics of Wind Turbines* (Routledge, 2015).

2. J. G. Schepers, "Engineering Models in Wind Energy Aerodynamics" (PhD thesis, Delft University of Technology, 2012), https://doi.org/10.4233/uuid:92123c07-cc12-4945-973f-103bd744ec87.

3. W. Yu, "The Wake of an Unsteady Actuator Disc" (PhD thesis, Delft University of Technology, 2018), https://doi.org/10.4233/uuid:0e3a2402-585c-41b1-81cf-a35753076dfc.

4. D. Micallef and T. Sant, "A Review of Wind Turbine Yaw Aerodynamics," in *Wind Turbines—Design, Control and Applications*, 1st ed., eds. A. G. Aissaoui and A. Tahour (IntechOpen, 2016), 27–53, https://doi.org/10.5772/63445.

5. C. Hur, T. Berdowski, C. S. Ferreira, K. Boorsma, and G. Schepers, "A Review of Momentum Models for the Actuator Disk in Yaw," in *AIAA Scitech 2019 Forum* (AIAA, 2019), https://doi.org/10.2514/6.2019-1799.

6. J. G. Leishman, "Challenges in Modelling the Unsteady Aerodynamics of Wind Turbines," *Wind Energy* 5, no. 2–3 (2002): 85–132, https://doi.org/10.1002/we.62.

7. W. Haans, T. Sant, G. van Kuik, and G. van Bussel, "Measurement of Tip Vortex Paths in the Wake of a HAWT Under Yawed Flow Conditions," *Journal of Solar Energy Engineering* 127, no. 4 (2005): 456–463, https://doi.org/10.1115/1.2037092.

8. J. Schepers, H. A. Madsen, T. Lutz, et al., *Final Report of IEA Task 29, Mexnet (Phase 1): Analysis of Mexico Wind Tunnel Measurements*." Tech. Rep. (International Energy Agency (IEA), 2012).

9. H. Glauert, *A General Theory of the Autogyro*." Tech. Rep. (HM Stationery Office, 1926).

10. R. P. Coleman, A. M. Feingold, and C. W. Stempin, *Evaluation of the Induced-Velocity Field of an Idealized Helicoptor Rotor*." Tech. Rep. (University of North Texas Libraries, UNT Digital Library, 1945).

11. S. Øye, "Induced Velocities for Rotors in Yaw," in *Proceedings of the Sixth IEA Symposium* (ECN, 1992).

12. G. A. M. van Kuik, "On the Generation of Vorticity by Force Fields in Rotor- and Actuator Flows," *Renewable Energy* 70 (2014): 124–128.

13. E. Branlard and M. Gaunaa, "Cylindrical Vortex Wake Model: Right Cylinder," *Wind Energy* 18, no. 11 (2015): 1973–1987, https://doi.org/10.1002/we.1800.

14. E. Branlard and M. Gaunaa, "Cylindrical Vortex Wake Model: Skewed Cylinder, Application to Yawed or Tilted Rotors," *Wind Energy* 19, no. 2 (2016): 345–358, https://doi.org/10.1002/we.1838.

15. E. Branlard, M. Gaunaa, and E. Machefaux, "Investigation of a New Model Accounting for Rotors of Finite Tip-Speed Ratio in Yaw or Tilt," *Journal of Physics: Conference Series* 524 (2014): 012124, https://doi.org/10.1088/1742-6596/524/1/012124.

16. E. Branlard and M. Gaunaa, "Superposition of Vortex Cylinders for Steady and Unsteady Simulation of Rotors of Finite Tip-Speed Ratio," *Wind Energy* 19, no. 7 (2016): 1307–1323, https://doi.org/10.1002/we.1899.

17. J. Schepers, "An Engineering Model for Yawed Conditions, Developed on Basis of Wind Tunnel Measurements," in *37th Aerospace Sciences Meeting and Exhibit* (AIAA, 1999), https://doi.org/10.2514/6.1999-39.

18. H. Rahimi, M. Hartvelt, J. Peinke, and J. G. Schepers, "Investigation of the Current Yaw Engineering Models for Simulation of Wind Turbines in BEM and Comparison With CFD and Experiment," *Journal of Physics: Conference Series* 753 (2016): 022016, https://doi.org/10.1088/1742-6596/753/2/022016.

19. H. Rahimi, A. Martinez Garcia, B. Stoevesandt, J. Peinke, and G. Schepers, "An Engineering Model for Wind Turbines Under Yawed Conditions Derived From High Fidelity Models," *Wind Energy* 21, no. 8 (2018): 618–633, https://doi.org/10.1002/we.2182.

20. J. R. Koza, "Genetic Programming: On the Programming of Computers by Means of Natural Selection," *Statistics and Computing* 4, no. 2 (1994): 87–112, https://doi.org/10.1007/BF00175355.

21. M. Schmidt and H. Lipson, "Distilling Free-Form Natural Laws From Experimental Data," *Science* 324, no. 5923 (2009): 81–85, https://doi.org/10.1126/science.1165893.

22. S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering Governing Equations From Data by Sparse Identification of Nonlinear Dynamical Systems," *Proceedings of the National Academy of Sciences* 113, no. 15 (2016): 3932–3937, https://doi.org/10.1073/pnas.1517384113.

23. M. Cranmer, "Interpretable Machine Learning for Science With PySR and SymbolicRegression.jl," preprint, arXiv, May 5, 2023, https://doi.org/10.48550/arXiv.2305.01582.

24. T. Mengel, P. Steffanic, C. Hughes, A. C. O. da Silva, and C. Nattrass, "Interpretable Machine Learning Methods Applied to Jet Background Subtraction in Heavy-Ion Collisions," *Physical Review C* 108, no. 2 (2023): L021901, https://doi.org/10.1103/PhysRevC.108.L021901.

25. V. Hruška, A. Furmanová, and M. Bednařík, "Analytical Formulae for Design of One-Dimensional Sonic Crystals With Smooth Geometry Based on Symbolic Regression," *Journal of Sound and Vibration* 597 (2025): 118821, https://doi.org/10.1016/j.jsv.2024.118821.

26. D. Wang, Y. Chen, and S. Chen, "Discovering an Interpretable Mathematical Expression for a Full Wind-Turbine Wake With Artificial Intelligence Enhanced Symbolic Regression," *Physics of Fluids* 36, no. 10 (2024): 105110, https://doi.org/10.1063/5.0221611.

27. P. Valsaraj, D. A. Thumba, K. Asokan, and K. S. Kumar, "Symbolic Regression-Based Improved Method for Wind Speed Extrapolation From Lower to Higher Altitudes for Wind Energy Applications," *Applied Energy* 260 (2020): 114270, https://doi.org/10.1016/j.apenergy.2019.114270.

28. D. Wang, D. Feng, K. Zhou, Y. Chen, S. Liao, and S. Chen, "Symbolic Regression-Enhanced Dynamic Wake Meandering: Fast and Physically Consistent Wind-Turbine Wake Modeling," preprint, arXiv, June 17, 2025, https://doi.org/10.48550/arXiv.2506.14403.

29. J. N. Sorensen and W. Z. Shen, "Numerical Modeling of Wind Turbine Wakes," *Journal of Fluids Engineering* 124, no. 2 (2002): 393–399, https://doi.org/10.1115/1.1471361.

30. J. Jonkman, S. Butterfield, W. Musial, and G. Scott, *Definition of a 5-MW Reference Wind Turbine for Offshore System Development*." Tech. Rep. (National Renewable Energy Laboratory (NREL), 2009).

31. Z. Yu, X. Zheng, and Q. Ma, "Study on Actuator Line Modeling of Two NREL 5-MW Wind Turbine Wakes," *Applied Sciences* 8, no. 3 (2018): 434, https://doi.org/10.3390/app8030434.

32. L. van den Bos, W. Bierbooms, A. Alexandre, B. Sanderse, and G. van Bussel, "Fatigue Design Load Calculations of the Offshore NREL 5 MW Benchmark Turbine Using Quadrature Rule Techniques," *Wind Energy* 23, no. 5 (2020): 1181–1195, https://doi.org/10.1002/we.2470.

33. OpenCFD Ltd., "OpenFOAM: User Guide v2306," (2023), https://develop.openfoam.com/Development/openfoam/-/tree/OpenFOAM-v2306.

34. P. Bachant, A. Goude, D. Mec, and M. Wosnik, *turbinesFoam/turbinesFoam: v0.1.1* (Zenodo, 2019), https://doi.org/10.5281/zenodo.3542301.

35. F. R. Menter, "Two-Equation Eddy-Viscosity Turbulence Models for Engineering Applications," *AIAA Journal* 32, no. 8 (1994): 1598–1605, https://doi.org/10.2514/3.12149.

36. J. Thé and H. Yu, "A Critical Review on the Simulations of Wind Turbine Aerodynamics Focusing on Hybrid RANS-LES Methods," *Energy* 138 (2017): 257–289, https://doi.org/10.1016/j.energy.2017.07.028.

37. Y. Li, W. Yu, A. Sciacchitano, and C. Ferreira, "Numerical Investigation of Regenerative Wind Farms Featuring Enhanced Vertical Energy Entrainment," *Wind Energy Science* 10, no. 4 (2025): 631–659, https://doi.org/10.5194/wes-10-631-2025.

38. J. Crank and P. Nicolson, "A Practical Method for Numerical Evaluation of Solutions of Partial Differential Equations of the Heat-Conduction Type," *Mathematical Proceedings of the Cambridge Philosophical Society* 43, no. 1 (1947): 50–67, https://doi.org/10.1017/S0305004100023197.

39. R. I. Issa, "Solution of the Implicitly Discretised Fluid Flow Equations by Operator-Splitting," *Journal of Computational Physics* 62, no. 1 (1986): 40–65, https://doi.org/10.1016/0021-9991(86)90099-9.

40. S. V. Patankar and D. B. Spalding, "A Calculation Procedure for Heat, Mass and Momentum Transfer in Three-Dimensional Parabolic Flows," *International Journal of Heat and Mass Transfer* 15, no. 10 (1972): 1787–1806, https://doi.org/10.1016/B978-0-08-030937-8.50013-1.

41. R. Poletto, T. Craft, and A. Revell, "A New Divergence Free Synthetic Eddy Method for the Reproduction of Inlet Flow Conditions for LES," *Flow, Turbulence and Combustion* 91, no. 3 (2013): 519–539, https://doi.org/10.1007/s10494-013-9488-2.

42. Y. Li, W. Yu, A. Sciacchitano, and C. Ferreira, "Wake Aerodynamic of Multi-Rotor System With Lifting-Devices Under Different Ambient Turbulence," *Journal of Physics: Conference Series* 3016 (2025): 012042, https://doi.org/10.1088/1742-6596/3016/1/012042.

43. Y. Li, W. Yu, and H. Sarlak, "Wake Structures and Performance of Wind Turbine Rotor With Harmonic Surging Motions Under Laminar and Turbulent Inflows," *Wind Energy* 27, no. 12 (2024): 1499–1525, https://doi.org/10.1002/we.2949.

44. N. Troldborg, "*Actuator Line Modeling of Wind Turbine Wakes*" (PhD thesis, Technical University of Denmark, 2009).

45. S. Naderi and F. Torabi, "Numerical Investigation of Wake Behind a HAWT Using Modified Actuator Disc Method," *Energy Conversion and Management* 148 (2017): 1346–1357, https://doi.org/10.1016/j.enconman.2017.07.003.

46. W. Z. Shen, R. Mikkelsen, J. N. Sørensen, and C. Bak, "Tip Loss Corrections for Wind Turbine Computations," *Wind Energy* 8, no. 4 (2005): 457–475, https://doi.org/10.1002/we.153.

47. L. A. Martínez-Tossas, M. J. Churchfield, and C. Meneveau, "A Highly Resolved Large-Eddy Simulation of a Wind Turbine Using an Actuator Line Model With Optimal Body Force Projection," *Journal of Physics: Conference Series* 2016; 753 082014, https://doi.org/10.1088/1742-6596/753/8/082014.

48. M. J. Churchfield, S. J. Schreck, L. A. Martinez, C. Meneveau, and P. R. Spalart, "An Advanced Actuator Line Method for Wind Energy Applications and Beyond," in *35th Wind Energy Symposium* (AIAA, 2017), https://doi.org/10.2514/6.2017-1998.

49. E. Gaertner, J. Rinker, L. Sethuraman, et al., *IEA Wind TCP Task 37: Definition of the IEA 15-Megawatt Offshore Reference Wind Turbine*." Tech. Rep. (National Renewable Energy Laboratory (NREL), 2020).

50. C. Bak, F. Zahle, R. Bitsche, et al., "The DTU 10-MW Reference Wind Turbine," in *Danish Wind Power Research* (Danmarks Tekniske Universitet, 2013).

51. M. Zormpa, Z. F. Arcos, X. Chen, C. R. Vogel, and R. H. Willden, "The Effect of Flow Sampling on the Robustness of the Actuator Line Method," *Wind Energy* 28, no. 1 (2025): e2965, https://doi.org/10.1002/we.2965.

52. P. Kidger, "On Neural Differential Equations," preprint, arXiv, February 4, 2022, https://doi.org/10.48550/arXiv.2202.02435.

53. S. Verstyuk and M. R. Douglas, "Machine Learning the Gravity Equation for International Trade," (2022), https://doi.org/10.2139/ssrn.4053795.

54. B. L. Davis and Z. Jin, "Discovery of a Planar Black Hole Mass Scaling Relation for Spiral Galaxies," *Astrophysical Journal Letters* 956, no. 1 (2023): L22, https://doi.org/10.3847/2041-8213/acfa98.

55. C. J. Soelistyo and A. R. Lowe, "Discovering Interpretable Models of Scientific Image Data With Deep Learning," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (IEEE, 2024), https://doi.org/10.1109/CVPRW63382.2024.00682.

56. G. F. Smits and M. Kotanchek, "Pareto-Front Exploitation in Symbolic Regression," in *Genetic Programming Theory and Practice II*, 1st ed., eds. U. M. O'Reilly, T. Yu, R. Riolo, and B. Worzel (Springer, 2005), 283–299.

57. Y. Li, H. Wang, Y. Li, et al., "Electron Transfer Rules of Minerals Under Pressure Informed by Machine Learning," *Nature Communications* 14, no. 1 (2023): 1815, https://doi.org/10.1038/s41467-023-37384-1.

58. K. I. Park and M. J. Park, *Fundamentals of Probability and Stochastic Processes With Applications to Communications* (Springer, 2018).

59. K. A. Kragh and M. H. Hansen, "Load Alleviation of Wind Turbines by Yaw Misalignment," *Wind Energy* 17, no. 7 (2014): 971–982, https://doi.org/10.1002/we.1612.

60. M. Bastankhah and F. Porté-Agel, "Effects of Positive Versus Negative Yaw Angles on Wind-Turbine Performance: An Application of BEM Theory," in *EGU General Assembly Conference Abstracts* (EGU General Assembly, 2014).

61. H. Sun, A. Sciacchitano, and W. Yu. "Supporting Data Belonging to Publication "An Engineering Model for Static Yawed Wind Turbines Based on Actuator Line Simulations and Symbolic Regression," 2025. 4TU Research Data, https://doi.org/10.4121/1258f8d9-a8a7-4912-9187-6bc35b217e19.

62. Delft High Performance Computing Centre (DHPC), "DelftBlue Supercomputer (Phase 2)," (2024), https://www.tudelft.nl/dhpc/ark:/44463/DelftBluePhase2.

63. H. A. Madsen, V. Riziotis, F. Zahle, et al., "Blade Element Momentum Modeling of Inflow With Shear in Comparison With Advanced Model Results," *Wind Energy* 15, no. 1 (2012): 63–81, https://doi.org/10.1002/we.493.

64. T. J. Larsen and A. M. Hansen, *How 2 HAWC2: The User's Manual* (Risø National Laboratory, 2007).

65. B. Wen, X. Tian, X. Dong, Z. Peng, W. Zhang, and K. Wei, "A Numerical Study on the Angle of Attack to the Blade of a Horizontal-Axis Offshore Floating Wind Turbine Under Static and Dynamic Yawed Conditions," *Energy* 168 (2019): 1138–1156, https://doi.org/10.1016/j.energy.2018.11.082.

66. ExaWind Community, "Actuator Line Model Calibration AMR Wind 0.1 Documentation," (2025), https://exawind.github.io/amr-wind/walkthrough/calibration.html.

Appendix A

Branlard's Yaw Engineering Model

Based on a skewed cylinder vortex wake model, Branlard's yaw engineering model is formulated using the Biot–Savart law applied to a complete vortex system. The tangential and normal induced velocities at the rotor disk are obtained by evaluating the contributions from each vortex element, and the total velocity field ($V_{ni,\text{total}}$ and $V_{ti,\text{total}}$) is computed by summing these components, as given in Equations (A1–A5). Among the vortex contributions, the tangential tip vortex is the most influential, followed by the root vortex and the longitudinal tip vortex [15].

$$V_{ni,total} = V_{ni,t} + V_{ni,r} + V_{ni,l}, \quad V_{ti,total} = V_{ti,t} + V_{ti,r} + V_{ti,l}, \tag{A1}$$

where t, r, and l denote the tangential tip, root, and longitudinal tip vortex components, respectively. Note that the induced velocity from the bound vortex at the rotor plane is zero and is therefore omitted.

Velocities induced by tangential tip vortex:

$$\begin{aligned} V_{ni,t}(r,\psi,\chi) &\approx V_{ni,0}\left[1+2F_t(r,\chi)\tan\left(\frac{\chi}{2}\right)\cos\psi_m\right], \\ V_{ti,t}(r,\psi,\chi) &\approx -\tan\left(\frac{\chi}{2}\right)\sin\psi_m\, V_{n,t}, \end{aligned} \tag{A2}$$

Velocities induced by root vortex:

$$V_{ni,r}(r,\theta,\chi) = \frac{\Gamma_r \sin\theta\sin\chi}{4\pi r(1-\cos\theta\sin\chi)}, V_{ti,r}(r,\theta,\chi) = \frac{\Gamma_r\cos\chi}{4\pi r(1-\cos\theta\sin\chi)} \tag{A3}$$

Velocities induced by longitudinal tip vortex:

$$\begin{aligned} V_{ni,l} &\approx -\gamma_l F_l(r)\tan^2\left(\frac{\chi}{2}\right)\sin(2\theta), \\ V_{ti,l} &\approx u_{y,0}\left(1+G_l(r)-F_l(r)\sec^2\left(\frac{\chi}{2}\right)\right)\cos\theta \end{aligned} \tag{A4}$$

$$\begin{aligned} G_l(\tilde{r}) &\approx 0.1(\tilde{r}+7\tilde{r}), \quad F_l(\tilde{r}) \approx 0.1(\tilde{r}-2\tilde{r}^3+6\tilde{r}^5), \quad u_{y,0}(x) \\ &= \gamma_l \frac{\sqrt{1+\tan^2\chi}-1}{2\tan\chi} \end{aligned} \tag{A5}$$

Appendix B

Code Verification

To verify the in-house BEM code and ALM results (turbinesFoam) used in this study, their predictions of normal and tangential induced velocities (V_{ni} and V_{ti}) and aerodynamic forces (f_n and f_t) are compared with results from several established numerical tools. The verification follows benchmark cases from Madsen et al. [63] and includes an actuator line model (ALM) [29] and a BEM implementations, HAWC2 [64]. The verification uses the NREL 5-MW wind turbine operating in uniform axial inflow, with a free-stream velocity of $V_\infty = 8\,\mathrm{m/s}$ at the optimal tip-speed ratio of $\lambda = 7.55$ [30]. Figure B1 shows the distributions of induced velocities and aerodynamic forces along radial positions. Overall, the in-house BEM framework and ALM results from this study shows strong agreement with the reference codes, although some differences are observed near the blade root and tip, primarily due to the inclusion of tip-loss corrections in BEM-based codes, which are absent in certain other models [65].

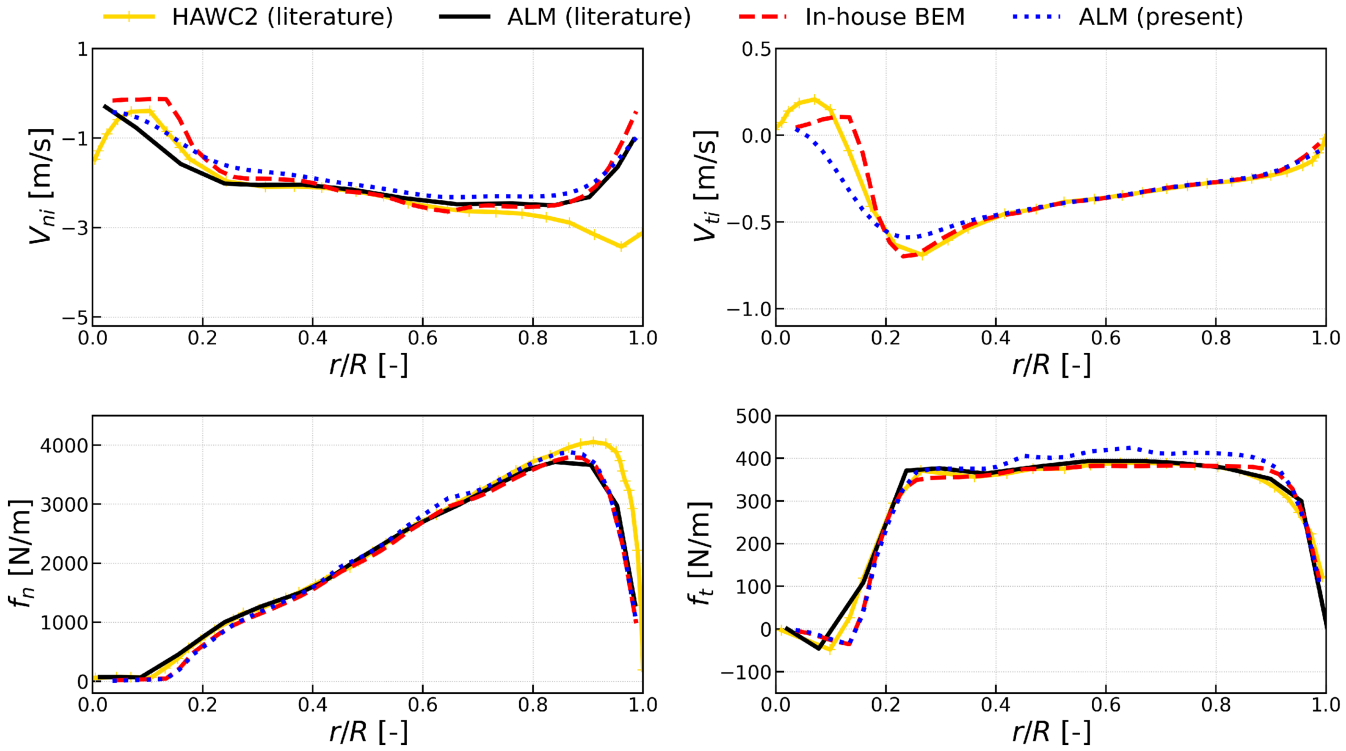

FIGURE B1 | Radial distributions of normal and tangential induced velocities (top left and right) and corresponding forces (bottom left and right) computed by different codes under the operating conditions of case Ue.1 in axial inflow.

Appendix C

Model Verification for the NREL 5-MW Turbine at $V_\infty = 11.4\ \mathrm{m/s}$ and $\lambda = 7.0$

The SR model is further evaluated under rated wind conditions at $V_\infty = 11.4$ m/s with a tip-speed ratio of $\lambda = 7.0$ and a yaw angle of $\gamma = 25°$ (case Ue.2). The ALM simulation is performed using the same setup as described in Section 2.1. Figures C1 and C2 show the azimuthal distributions of induced velocity and aerodynamic forces from BEM results incorporating the three models discussed previously, compared against ALM results. The results indicate that similar trends are observed as in the below-rated case Ue.1. Above midspan, differences among the models are minor, while from midspan toward the root, the SR model more accurately captures the phase and better predicts the location of maximum forces.

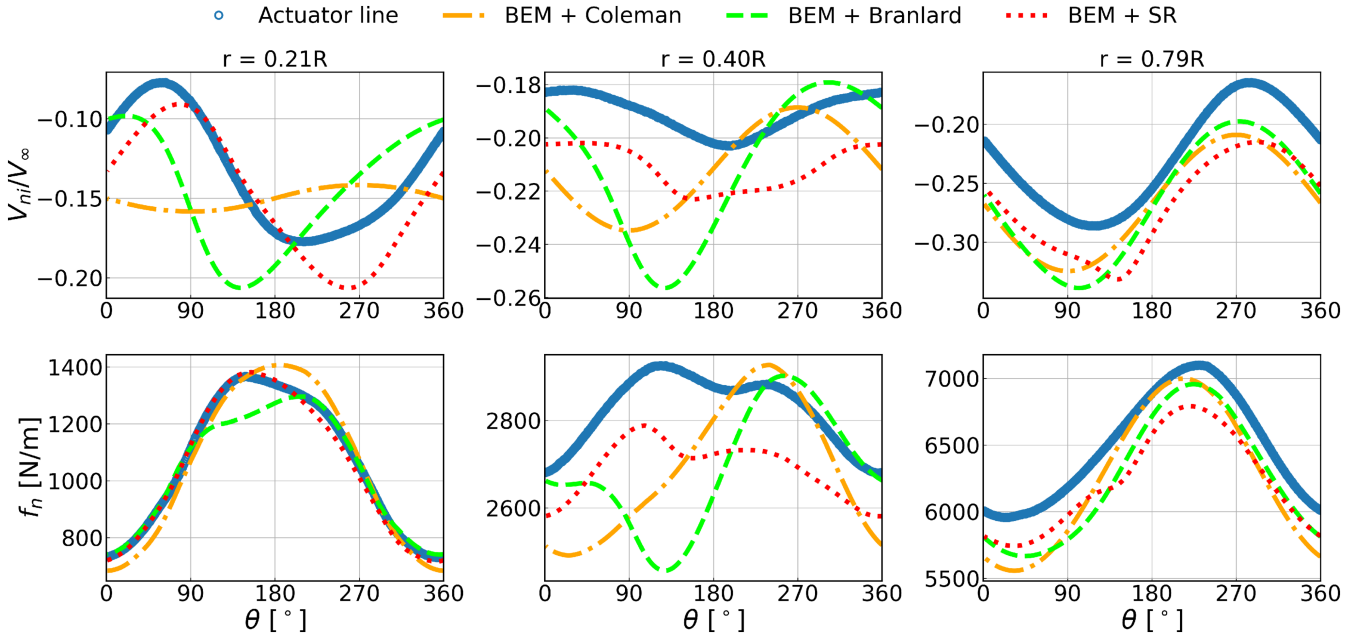

FIGURE C1 | Azimuthal distributions of normal induced velocity V_{ni} and normal force f_n at three radial positions for $\gamma = 25°$. Each subplot represents a radial location, with ALM results as reference and overlaid predictions from BEM simulations using the Coleman, Branlard, and SR models for comparison.

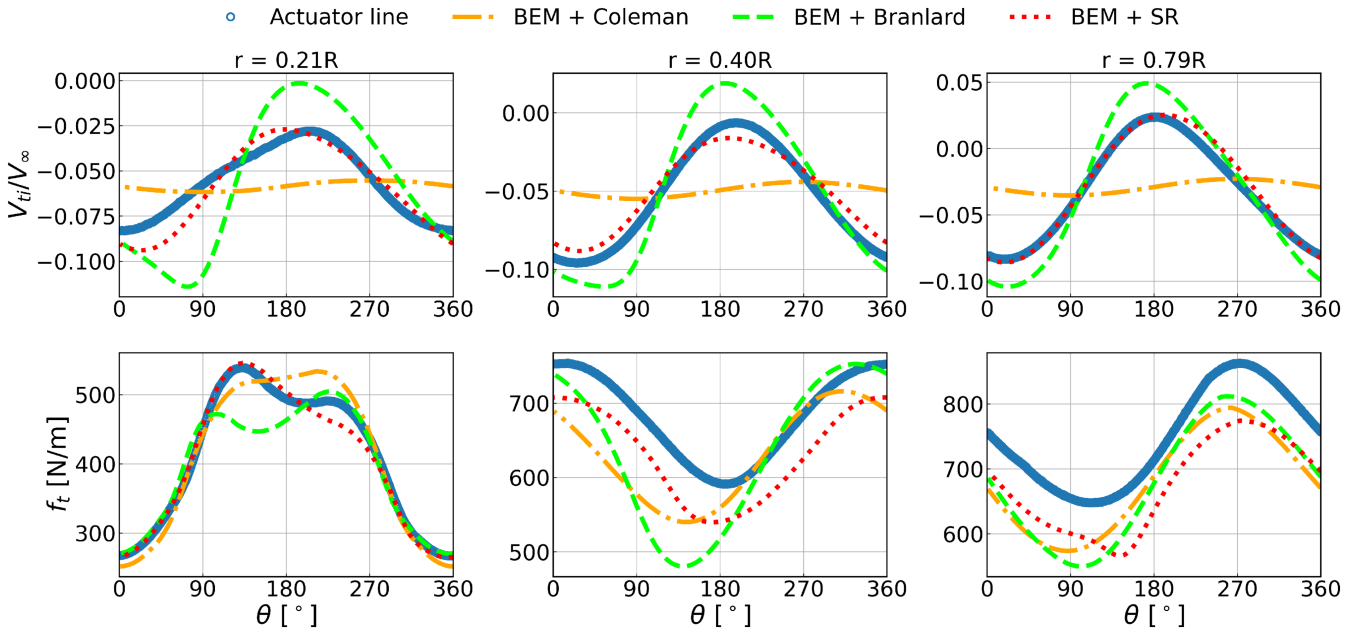

FIGURE C2 | Azimuthal distributions of normal induced velocity V_{ti} and tangential force f_t at three radial positions for $\gamma = 25°$. Each subplot represents a radial location, with ALM results as reference and overlaid predictions from BEM simulations using the Coleman, Branlard, and SR models for comparison.

Appendix D

Model Verification for the IEA 15-MW Turbine at $V_\infty = 9.0\ \mathrm{m/s}$ and $\lambda = 9.0$

Figures D1 and D2 illustrate the azimuthal distributions of induced velocities and aerodynamic forces for an entirely unseen wind turbine model the IEA 15-MW reference turbine. The inflow conditions correspond to $V_\infty = 9.0$ m/s and a tip-speed ratio of $\lambda = 9.0$ at a yaw angle of $\gamma = 25°$ (case Ue.3). The reference ALM simulation utilized here is based on the configuration documented in [66] and is not reiterated in detail. Consistent with the setup in the current study, this ALM simulation excludes complex physical factors such as tower influence, aeroelastic effects and blade prebending. It is demonstrably clear that the SR model provides superior agreement with the ALM results in terms of both amplitude and phase, a performance that is maintained even when applied to entirely unseen data and a new turbine configuration.

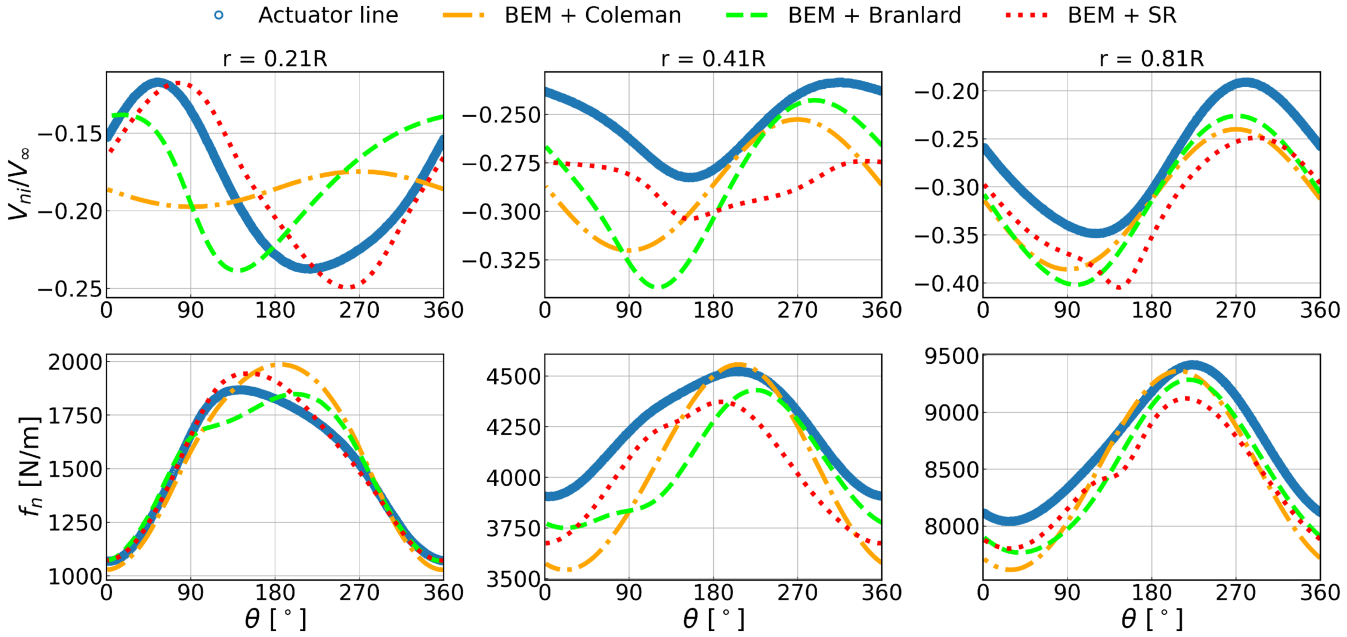

FIGURE D1 | Azimuthal distributions of normal induced velocity V_{ni} and normal force f_n at three radial positions for $\gamma = 25°$. Each subplot represents a radial location, with ALM results as reference and overlaid predictions from BEM simulations using the Coleman, Branlard, and SR models for comparison.

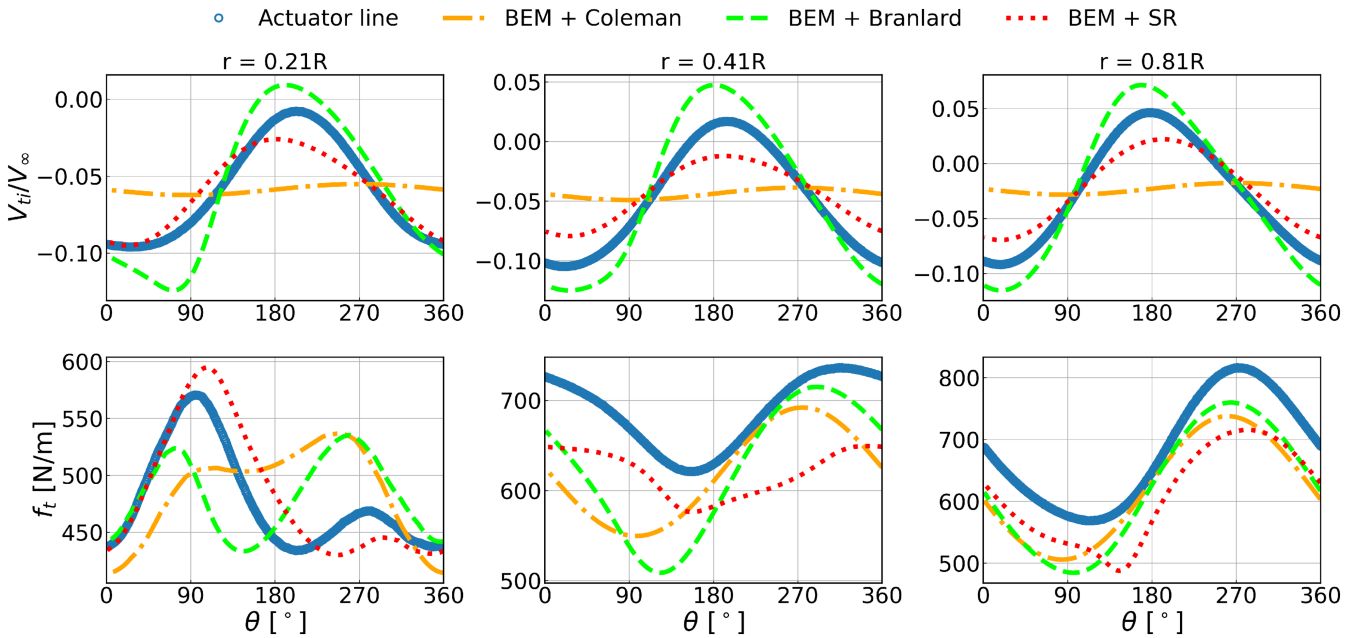

FIGURE D2 | Azimuthal distributions of normal induced velocity V_{ti} and tangential force f_t at three radial positions for $\gamma = 25°$. Each subplot represents a radial location, with ALM results as reference and overlaid predictions from BEM simulations using the Coleman, Branlard, and SR models for comparison.