

An Online Prediction Method for Transient Frequency Response in New Energy Grids Based on Deep Integration of WAMS Data and Physical Model

Original

An Online Prediction Method for Transient Frequency Response in New Energy Grids Based on Deep Integration of WAMS Data and Physical Model / Yan, K., Hu, Y., Xu, H., Huang, T., Long, Y., Wang, T.. - In: ENTROPY. - ISSN 1099-4300. - 27:11(2025). [10.3390/e27111145]

Availability:

This version is available at: 11583/3009208 since: 2026-03-25T10:15:16Z

Publisher:

Multidisciplinary Digital Publishing Institute (MDPI)

Published

DOI:10.3390/e27111145

Terms of use:







This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

Article

An Online Prediction Method for Transient Frequency Response in New Energy Grids Based on Deep Integration of WAMS Data and Physical Model

Kailin Yan ¹, Yi Hu ^{1,*}, Han Xu ¹, Tao Huang ², Yang Long ³ and Tao Wang ¹

¹ School of Electrical Engineering and Electronic Information, Xihua University, Chengdu 610039, China; yankailin@stu.xhu.edu.cn (K.Y.); xuduidui@stu.xhu.edu.cn (H.X.); wangatao@163.com (T.W.)

² Department of Energy, Politecnico di Torino, 10129 Turin, Italy; tao.huang@polito.it

³ School of Electrical Engineering, Southwest Jiaotong University, Chengdu 611756, China; zhouchuanjiang@stu.xhu.edu.cn

* Correspondence: 1220200001@xhu.edu.cn; Tel.: +86-18482136949

Abstract

The integration of a high proportion of renewable energy has significantly reduced the grid inertia level and markedly increased the risk of transient frequency instability in power systems. Meanwhile, the large-scale integration of diverse heterogeneous resources—such as wind power, photovoltaics, energy storage, and high voltage direct current (HVDC) transmission systems—has considerably enriched the portfolio of frequency regulation assets in modern power grids. However, the marked disparities in the dynamic response characteristics and actuation speeds among these resources introduce significant nonlinearity and high-dimensional complexity into the system's transient frequency behavior. As a result, conventional methods face considerable challenges in achieving accurate and timely prediction of such responses. However, the substantial differences in the frequency regulation characteristics and response speeds of these resources have led to a highly nonlinear and high-dimensional complex transient frequency response process, which is difficult to accurately and rapidly predict using traditional methods. To address this challenge, this paper proposes an online prediction method for transient frequency response that deeply integrates physical principles with data-driven approaches. First, a frequency dynamic response analysis model incorporating the frequency regulation characteristics of multiple resource types is constructed based on the Single-Machine Equivalent (SME) method, which is used to extract key features of the post-fault transient frequency response. Subsequently, information entropy theory is introduced to quantify the informational contribution of each physical feature, enabling the adaptive weighted fusion of physical frequency response features and Wide-Area Measurement System (WAMS) data. Finally, a physics-guided machine learning framework is proposed, in which the weighted physical features and the complete frequency curve predicted by the physical model are jointly embedded into the prediction process. An MLP-GRU-Attention model is designed as the data-driven predictor for frequency response. A physical consistency constraint is incorporated into the loss function to ensure that predictions strictly adhere to physical laws, thereby enhancing the accuracy and reliability of the transient frequency prediction model. Case studies based on the modified IEEE 39-bus system demonstrate that the proposed method significantly outperforms traditional data-driven approaches in terms of prediction accuracy, generalization capability under small-sample conditions, and noise immunity. This provides a new avenue for online frequency security awareness in renewable-integrated power systems with multiple heterogeneous frequency regulation resources.



check for updates

Academic Editor:

Ali Mohammad-Djafari

Received: 17 October 2025

Revised: 2 November 2025

Accepted: 5 November 2025

Published: 10 November 2025

Citation: Yan, K.; Hu, Y.; Xu, H.; Huang, T.; Long, Y.; Wang, T. An Online Prediction Method for Transient Frequency Response in New Energy Grids Based on Deep Integration of WAMS Data and Physical Model. *Entropy* **2025**, *27*, 1145. <https://doi.org/10.3390/e27111145>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: new energy grid; transient frequency prediction; WAMS data; information entropy; physics-guided learning; MLP-GRU-Attention model

1. Introduction

In recent years, the penetration of renewable energy generation has continuously increased, while the proportion of conventional synchronous generators in the installed capacity continues to decline. Renewable energy generation units, such as wind and photovoltaic power systems, are typically integrated via power electronic interfaces. Their output power is decoupled from the system frequency, thus unable to provide inherent inertial support to the grid [1–3]. With the continuous increase in the penetration of renewable energy generation, the system's equivalent inertia has significantly decreased, resulting in a higher Rate of Change of Frequency (RoCoF). During power disturbances, this leads to more severe transient frequency fluctuations, which increases the risk of triggering security control measures such as Under-Frequency Load Shedding (UFLS), thereby seriously jeopardizing the secure and stable operation of the power system [4–6].

Meanwhile, to mitigate the inherent variability and uncertainty of renewable energy and enhance renewable energy integration, new power electronic devices such as energy storage systems and HVDC transmission have been deployed on a large scale in power grids [7,8]. As a result, modern renewable-rich power systems incorporate various heterogeneous frequency regulation resources, including wind power, photovoltaics, energy storage, and HVDC technologies. While these resources enhance the system's regulatory capacity, their frequency regulation characteristics, such as response speed, duration, and actuation mechanisms, differ significantly [9,10]. The coordination and cooperation among various resources are highly complex, resulting in a transient frequency response process characterized by strong nonlinearity, high-dimensional complexity, and significant uncertainty, thereby further increasing the difficulty of accurately predicting transient frequency.

Against this backdrop, achieving fast and accurate online prediction of transient frequency response has become a critical prerequisite for deploying emergency control strategies and ensuring the secure and stable operation of power grids [11]. Currently, methods for analyzing post-disturbance frequency responses in power systems can be broadly categorized into physics-based models and data-driven approaches [12]. Physics-based modeling methods, such as time-domain simulation [13] and equivalent model-based approaches [14], offer clear physical mechanisms, but these methods exhibit inherent limitations: time-domain simulation is computationally intensive and often unsuitable for online applications; although equivalent model-based approaches improve efficiency through simplification, they struggle to accurately capture the complex dynamic interactions among multiple heterogeneous resources, resulting in limited model accuracy.

Data-driven approaches offer novel solutions to the aforementioned challenges [15–21]. Reference [15] proposed a method for predicting power system frequency response (SFR) patterns after large disturbances based on a Deep Belief Network (DBN), enabling rapid frequency security assessment. Reference [16] developed an SFR model using deep transfer learning for wind-thermal hybrid power systems and obtained key parameters of the SFR model via an improved Recurrent Neural Network (RNN) to achieve frequency response prediction based on an equivalent model method. Reference [17] introduced a Frequency Security Predictor (FSPM) based on a one-dimensional Convolutional Neural Network (1D-CNN) and a temporal-feature attention module (TFAM), enabling simultaneous prediction of both Frequency Danger Level (FDL) and Time Security Margin (TSM). Reference [18] employed an attention-based Long Short-Term Memory (LSTM) neural network to predict

grid frequency curves and used the Whale Optimization Algorithm (WOA) to optimize network parameters for improved prediction accuracy. Reference [19] applied a pruned one-dimensional time-aware CNN to analyze power system frequency stability and determine the optimal frequency regulation cost. Reference [20] implemented real-time grid frequency monitoring using a machine learning technique combining Nonlinear Auto Regressive (NAR) neural networks and LSTM networks based on wide-area measurement data. However, these methods heavily rely on large amounts of high-quality training data, exhibit “black-box” characteristics, suffer from limited generalizability and interpretability, may produce predictions that violate physical laws, and lack reliability in extreme disturbance scenarios with scarce data.

Given the respective limitations of physical models and data-driven methods, integrating the two approaches to leverage their respective strengths has emerged as a promising pathway to overcome current technical bottlenecks [22–24]. Based on real-time WAMS data, this paper proposes an online prediction method for transient frequency response that deeply integrates information entropy weighting and physics-guided learning. The main contributions of this paper are as follows:

- (1) A physics-based frequency response analysis model incorporating the frequency regulation characteristics of multiple heterogeneous resources is developed to provide critical physical features for data-driven approaches.
- (2) Information entropy theory is introduced to perform quantitative assessment and adaptive weighting of physical features, thereby enhancing the contribution of critical features in the model input.
- (3) A physics-guided machine learning framework is proposed, which integrates the weighted physical features along with the complete frequency curve predicted by the physical model into the prediction process.
- (4) An MLP-GRU-Attention hybrid model is designed as the data-driven engine, and a physical consistency constraint is incorporated into the loss function, thereby ensuring that the prediction results strictly adhere to physical laws.

The remainder of this paper is organized as follows: Section 2 establishes a physics-based frequency response model for renewable-integrated power systems and derives the calculation method for key features. Section 3 elaborates on the architecture of the prediction model, which integrates information entropy weighting and a physics-guided MLP-GRU-Attention framework. Section 4 introduces the offline training and online prediction workflow of the proposed method. Section 5 validates the superiority of the approach through case studies on a modified IEEE 39-bus system. Section 6 concludes the paper.

2. Physical Model Analysis of Dynamic Frequency Response in New Energy Grids

2.1. New Energy Grid Frequency Response Model

Focusing on a typical renewable-energy integrated power system as illustrated in Figure 1, this paper takes into account the combined influence of conventional thermal units, renewable energy sources, energy storage systems, HVDC transmission, and load demand on frequency dynamics. An aggregate modeling approach is adopted to develop a physical frequency response model that incorporates multi-resource frequency regulation [14,25]. The dynamic responses of heterogeneous frequency regulation resources—including wind power, photovoltaic systems, energy storage, and HVDC transmission—are aggregated and represented by a low-order equivalent model. The frequency regulation control parameters of this equivalent model can be identified from actual measurement data.

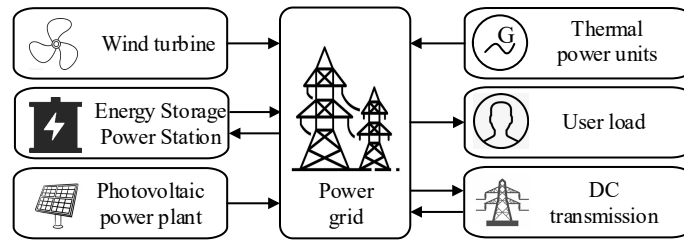


Figure 1. Schematic diagram of new energy grid.

(1) Frequency Response Model for Conventional Thermal Power Units

The frequency response characteristics of conventional thermal power units are primarily determined by the swing equation of synchronous generators and the dynamic characteristics of their governor systems [26]. Following a power disturbance after a system fault, the swing equation of the synchronous generator is described by

$$2H \frac{d\Delta f}{dt} = \Delta P_m - \Delta P_e - D\Delta f \tag{1}$$

where H denotes the inertia time constant; ΔP_m and ΔP_e represent the variations in mechanical power and electromagnetic power, respectively; and D is the damping coefficient.

At the instant of disturbance, the change in electromagnetic power ΔP_e primarily originates from the external disturbance power ΔP_d , which can be expressed as

$$\Delta P_e = \Delta P_d \tag{2}$$

During the primary frequency regulation process, the governor adjusts the mechanical power through droop control:

$$\Delta P_m = -K_G \Delta f \tag{3}$$

where K_G is the inverse of the generator’s droop coefficient.

Substituting Equations (2) and (3) into Equation (1) yields the system frequency response equation:

$$\frac{2H\Delta f}{dt} + (K_G + D)\Delta f = -\Delta P_d \tag{4}$$

Based on the equations presented above, the frequency response transfer function for the single-machine system can be derived as follows:

$$\frac{\Delta f(s)}{\Delta P_d(s)} = -\frac{1}{2Hs + D + K_G} \tag{5}$$

The frequency regulation model of a conventional thermal power unit in the power system is illustrated in Figure 2. To improve computational efficiency and meet the real-time requirements for online frequency prediction, the original model, which includes components such as reheat turbines, has been appropriately simplified while preserving its essential dynamic characteristics. Certain nonlinearities and slow dynamic processes are neglected, with only the key primary frequency regulation and inertial response features retained.

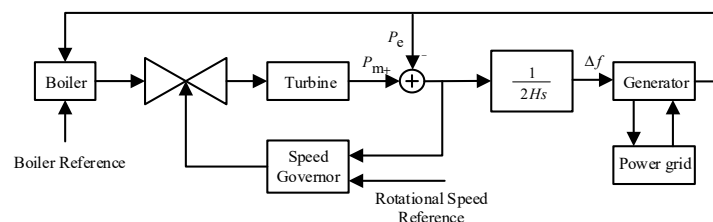


Figure 2. Frequency regulation model of conventional thermal power units.

For practical multi-machine systems, the process of synchronous oscillation power exchange between units is neglected, and an equivalent aggregate unit is employed to represent the system’s average frequency response. By aggregating the swing equations and governor dynamics of all units into a single unified model, the overall representation focuses solely on power-frequency range responses, while neglecting slower dynamics such as boiler systems. This leads to the frequency response model of a multi-thermal-unit power grid, as illustrated in Figure 3.

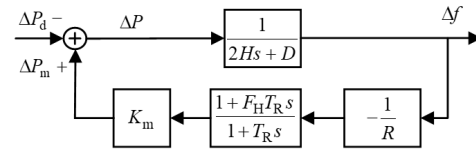


Figure 3. Single-machine equivalent frequency response model of the power system, where ΔP_d denotes the external disturbance power; ΔP_m represents the primary frequency regulation power of the synchronous generators; ΔP corresponds to the system accelerating power; K_m is the gain coefficient; F_H indicates the ratio of high-pressure turbine output power to the total mechanical power; and T_R signifies the reheat time constant.

(2) Frequency Response Characteristics of Wind Turbine Generators

Currently, grid-connected wind turbines widely employ grid-following converters, which inherently lack the ability to autonomously establish grid voltage and frequency and cannot provide natural inertial support like synchronous machines. To address this, frequency-regulating wind turbines typically adopt synthetic inertia control strategies that emulate the frequency regulation behavior of synchronous generators. This is achieved by incorporating virtual inertia control and primary frequency regulation into the wind turbine control system, including the release of rotational kinetic energy stored in the wind turbine rotor to provide inertial support, and the utilization of reserve capacity to deliver primary frequency regulation [27]. The virtual inertial response power of wind turbines can be described as follows:

$$P_{wd} = -k_d \frac{d\Delta f}{dt} \tag{6}$$

where k_d is the virtual inertia time constant and Δf is the system frequency deviation.

The power increment of wind turbine units participating in primary frequency regulation can be represented as follows:

$$P_{wp} = -k_p \Delta f \tag{7}$$

where k_p is the proportional parameter.

Therefore, the synthetic inertia control strategy model for frequency-regulating wind turbines is developed, as illustrated in Figure 4.

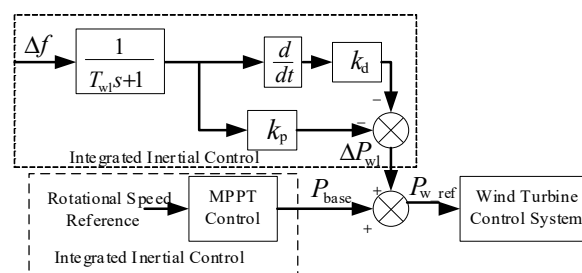


Figure 4. Frequency response model of wind turbine generator system.

In Figure 4, point P_{base} is set at the deloading operating point based on the system’s frequency regulation demand.

As can be seen from Figure 4, the output power ΔP_{wl} of the synthetic inertia control strategy, accounting for the response time of wind turbine converters, can be expressed as follows:

$$\Delta P_{wl} = P_{wd} + P_{wp} = -\frac{k_d s + k_p}{T_{wl} s + 1} \Delta f \tag{8}$$

where T_{wl} denotes the frequency regulation time constant of the grid-following wind turbine, Δf represents the system frequency deviation, and s is the Laplace operator.

Therefore, the reference operating power P_{w_ref} of a frequency-regulating wind turbine is obtained by combining the base operating power P_{base} during deloading operation with the additional frequency regulation power ΔP_{wl} generated by virtual inertia and primary frequency regulation control, expressed as follows:

$$P_{w_ref} = P_{base} + \Delta P_{wl} \tag{9}$$

In the equation, P_{base} represents the actual power setpoint of the wind turbine under deloaded operation, which is lower than the maximum capture power P_{MPPT} at the current wind speed. This ensures the wind turbine maintains upward reserve capacity. ΔP_{wl} denotes the additional frequency regulation power, which includes both the reserve capacity activated by the wind turbine’s participation in primary frequency regulation control and the power sourced from the release of rotor kinetic energy through virtual inertial control, which refers to the conversion of kinetic energy into electrical energy by temporarily reducing rotor speed to provide rapid power support. To ensure that the wind turbine operates within its physical limits, the total output power $P_{base} + \Delta P_{wl}$ must always remain less than P_{MPPT} .

(3) Frequency Response Characteristics of Photovoltaic (PV) and Energy Storage Systems

PV power plants typically participate in grid frequency regulation through deloading operation or integrated energy storage systems. Given the inherent volatility and unpredictability of PV power generation, its rapid frequency response characteristics are comparable to those of energy storage systems. Therefore, this paper models the frequency regulation contributions of PV power plants and their associated energy storage systems in a unified manner [28,29]. Both PV plants and energy storage systems, which are typically grid-connected via power electronic converters, employ droop control strategies to provide primary frequency regulation. Their frequency response model is illustrated in Figure 5.

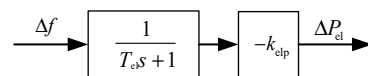


Figure 5. Frequency response model of PV and energy storage system.

Therefore, the auxiliary frequency regulation output power ΔP_{el} from the PV plant and the energy storage system can be expressed as follows:

$$\Delta P_{el} = -\frac{k_{elp}}{T_{el} s + 1} \Delta f \tag{10}$$

where k_{elp} denotes the frequency droop control coefficient of the energy storage system, and T_{el} represents its control time constant.

(4) Frequency Response Characteristics of DC Transmission Systems

The HVDC transmission system, equipped with a Frequency Limiting Controller (FLC), can provide short-term transient frequency support for renewable-energy-integrated power grids, representing a form of transient ancillary service. The design objective of the FLC is to rapidly inject active power during the initial stage of a system disturbance when significant frequency deviation occurs, thereby restraining frequency changes. Its control

strategy is centered on detecting when the system frequency deviation exceeds a preset dead band, upon which a proportional-integral (PI) controller is employed to generate an additional power command, dynamically adjusting the active power reference of the DC link [30].

The FLC is not suitable for long-term operation under off-nominal DC voltage. To avoid conflicts with DC system protection settings and ensure long-term grid stability, the FLC typically adopts a reverse reset strategy. This strategy is activated once the frequency deviation returns within the dead band, gradually resetting the additional power reference to zero and restoring the DC system to its rated operating state. This mechanism effectively mitigates risks such as equipment overload, voltage limit violations, and maloperation of protection relays. The control structure is illustrated in Figure 6.

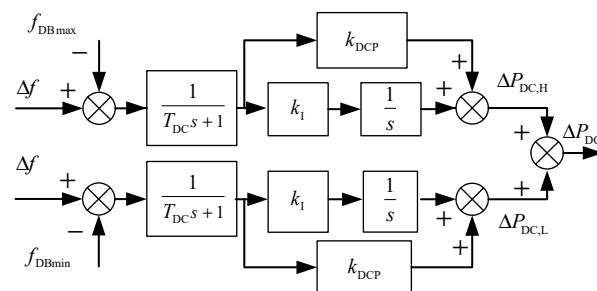


Figure 6. DC FLC frequency response model.

When the frequency deviation exceeds the control deadband, the positive and negative additional auxiliary power $\Delta P_{DC,H}$ and $\Delta P_{DC,L}$, as well as the total additional control regulatory signal ΔP_{DC} of the HVDC transmission system, are given by the following:

$$\begin{cases} \Delta P_{DC,H} = k_{DCP}(\Delta f - f_{DBmax}) + \int k_I(\Delta f - f_{DBmax})dt \\ \Delta P_{DC,L} = k_{DCP}(\Delta f - f_{DBmin}) + \int k_I(\Delta f - f_{DBmin})dt \\ \Delta P_{DC} = \Delta P_{DC,H} + \Delta P_{DC,L} \end{cases} \quad (11)$$

where f_{DBmax} and f_{DBmin} represent the forward and reverse dead band values of the FLC control; k_{DCP} and k_I denote the proportional and integral coefficients of the PI controller, respectively; and T_{DC} is the DC frequency regulation time constant. The proportional term $k_{DCP}(\Delta f - \Delta f_{DBmax})$ and the integral term $\int k_I(\Delta f - f_{DBmax})dt$ in the equation correspond to the proportional control and integral control of the PI controller in Figure 6, respectively. As long as the frequency deviation is not completely eliminated, the integral term will continue to accumulate over time, thereby continuously adjusting the DC power reference value. This mechanism aims to ultimately eliminate steady-state frequency errors and ensure the accuracy of frequency recovery.

(5) Multi-Resource Frequency Response Model

Based on the frequency regulation models of the various resources described above, the corresponding transfer functions characterizing their frequency regulation properties are derived and summarized in Table 1.

Based on Table 1, the integrated frequency response model of the renewable-energy-integrated power system is constructed using the transfer function aggregation method, with its structure shown in Figure 7 [31]. The first branch represents the equivalent transfer function of the system rotor motion equation, characterizing the overall inertial response and damping characteristics of the system. The subsequent four branches correspond to the transfer function models of the synchronous generator governor, wind turbine virtual inertial control, photovoltaic and energy storage droop control, and HVDC FLC, respectively.

Table 1. Frequency response transfer function of various FM resources.

Control Measures	Auxiliary Power Transfer Function
Synchronous Generator Governor	$G_m = \frac{\Delta P_m}{\Delta f_i} = \frac{K_m(1+F_H T_R s)}{R(1+T_R s)}$
Virtual Inertia Control for Fans	$G_{wl} = \frac{\Delta P_{wl}}{\Delta f_i} = \frac{k_d s + k_p}{T_{wl} s + 1}$
Photovoltaic and Energy Storage Droop Control	$G_{el} = \frac{\Delta P_{el}}{\Delta f_i} = \frac{k_{elp}}{T_{el} s + 1}$
DC Transmission FLC	$G_{DC} = \frac{\Delta P_{DC}}{\Delta f_i} = \frac{k_{DCP} s + k_I}{T_{DC} s^2 + s}$

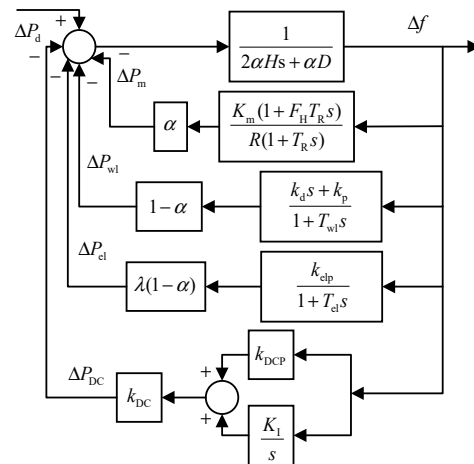


Figure 7. Frequency response model of new energy grid considering the participation of multiple frequency regulation resources.

Based on this framework, the system swing equation can be written as follows:

$$(2\alpha Hs + \alpha D)\Delta f = \Delta P_d - \Delta P_m - \Delta P_{wl} - \Delta P_{el} - \Delta P_{DC} \tag{12}$$

Accordingly, the Laplace-domain expression of the frequency response in the renewable-integrated power system is derived based on Table 1 as follows:

$$\Delta f = \frac{\Delta P_d}{2\alpha H + \alpha D + \frac{\alpha K_m(1+F_H T_R s)}{R(1+T_R s)} + (1-\alpha)\frac{k_d s + k_p}{1+T_{wl} s} + \frac{(1-\alpha)\lambda k_{elp}}{1+T_{el} s} + k_{DC}(k_{DCP} + \frac{K_I}{s})} \tag{13}$$

where α denotes the proportion of synchronous generator capacity to the total frequency regulation resource capacity in the system, quantifying the contribution weight of synchronous units in the frequency response; λ represents the ratio of the capacity of the energy storage system coupled with wind turbines to the total wind power capacity, reflecting the enhancement degree of energy storage to the frequency regulation capability of wind turbines. Both are capacity proportion coefficients determined based on the actual configuration of the system’s frequency regulation resources, and their specific values can be obtained through statistical calculation of system parameters and resource capacity.

$$\Delta f_{eq} \approx \frac{R_{eq}}{D_{eq} R_{eq} + K_{meq}} \frac{(1 + T_{Req} s)\Delta P_{deq}}{s(s^2 + 2\zeta\omega_n + \omega_n^2)} \tag{14}$$

$$\omega_n^2 = \frac{D_{eq} R + K_m}{2H_{eq} R T_R} \tag{15}$$

$$\zeta = \omega_n \left(\frac{2H_{eq}R + D_{eq}RT_R + K_m F_H T_R}{2(D_{eq}R + K_m)} \right) \tag{16}$$

$$H_{eq} = 2\alpha H + (1 - \alpha)k_d \tag{17}$$

$$D_{eq} = \alpha D + (1 - \alpha)k_p + (1 - \alpha)\lambda k_{elp} + k_{DC}k_{DCP} \tag{18}$$

Then, by applying the inverse Laplace transform, the time-domain analytical solution of the frequency response in the renewable-integrated power system incorporating multi-resource frequency regulation is obtained as follows:

$$\Delta f_{eq}(t) = \frac{R\Delta P_d}{D_{eq}R + K_m} [1 - e^{-\zeta\omega_n t} \cos(\sqrt{1 - \zeta^2}\omega_n t) - \frac{\zeta - \omega_n T_R}{\sqrt{1 - \zeta^2}} e^{-\zeta\omega_n t} \sin(\sqrt{1 - \zeta^2}\omega_n t)] \tag{19}$$

2.2. Key Characteristics of Frequency Response in New Energy Grids

Based on the time-domain analytical solution of the SFR described in Equation (17), five key transient frequency response features with clear physical interpretations and computational tractability can be derived. These features concisely yet comprehensively characterize the frequency dynamics following a disturbance, and serve as essential inputs for the subsequent deep integration of WAMS data and physical knowledge. They include, but are not limited to, the following:

(1) Rate of Change of Frequency, RoCoF Δf_{begin}

This characteristic reflects the initial rate of frequency decline following a disturbance and serves as a critical indicator for assessing the system inertia level. By taking the partial derivative of Equation (17) with respect to time, the RoCoF can be obtained as follows:

$$\Delta f_b = \frac{d\Delta f_{eq}(t)}{dt} = \frac{R\Delta P_d}{D_{eq}R + K_m} \omega_n e^{-\zeta\omega_n t} \left[\frac{1 - \zeta\omega_n T_R}{\sqrt{1 - \zeta^2}} \sin(\sqrt{1 - \zeta^2}\omega_n t) + \omega_n T_R \cos(\sqrt{1 - \zeta^2}\omega_n t) \right] \tag{20}$$

Evaluating at time $t = 0$ yields the initial RoCoF, denoted as Δf_{begin} :

$$\Delta f_{begin} = \left. \frac{d\Delta f(t)}{dt_{con}} \right|_{t=0} = \frac{R\Delta P_d}{D_{eq}R + K_m} \sin \varphi \omega_n T_R \tag{21}$$

(2) Minimum Time to Arrival t_{nadir}

This characteristic represents the time taken for the frequency to decline to its nadir, reflecting the response speed of the system's frequency regulation resources. The maximum frequency deviation occurs at time t_{nadir} . Given that the frequency nadir necessarily corresponds to an extremum point, where the partial derivative of Equation (17) with respect to time equals zero (i.e., the moment corresponding to $d\Delta f_{eq}/dt = 0$), the expression for t_{nadir} can be derived as follows:

$$t_{nadir} = \frac{1}{\sqrt{1 - \zeta^2}\omega_n} \arctan \left(\frac{\sqrt{1 - \zeta^2}\omega_n T_R}{\zeta\omega_n T_R - 1} \right) \tag{22}$$

(3) Lowest frequency Δf_{min}

This characteristic corresponds to the lowest point (nadir) reached during the frequency response process, which directly determines whether safety control measures such as UFLS will be triggered. Based on the provided analytical expression of the frequency re-

sponse and the expression for the time of maximum frequency deviation t_{nadir} , substituting t_{nadir} into the frequency response equation yields the maximum frequency deviation value:

$$\Delta f_{\min} = \frac{R\Delta P_d}{D_{\text{eq}}R + K_m} [1 - e^{-\xi\omega_n t_{\text{nadir}}} \cos(\sqrt{1 - \xi^2}\omega_n t_{\text{nadir}}) - \frac{\xi - \omega_n T_R}{\sqrt{1 - \xi^2}} e^{-\xi\omega_n t_{\text{nadir}}} \sin(\sqrt{1 - \xi^2}\omega_n t_{\text{nadir}})] \quad (23)$$

(4) Steady-state frequency f_{ss}

This characteristic signifies the frequency value at which the system settles into a new stable operating state following a disturbance, reflecting the capacity and droop characteristics of the frequency regulation resources. By taking the limit as t_{con} approaches infinity, the steady-state frequency f_{sta} can be calculated as follows:

$$f_{\text{ss}} = \lim_{t \rightarrow \infty} \Delta f_{\text{eq}}(t) = \frac{R\Delta P_d}{D_{\text{eq}}R + K_m} \quad (24)$$

(5) Steady-State Frequency Settling Time t_{ss}

This characteristic is defined as the time required for the SFR to enter and remain within a $\pm 2\%$ error band of the steady-state value. It reflects the speed of frequency recovery and can be expressed as follows:

$$t_{\text{ss}} = -\frac{1}{\xi\omega_n} \ln\left(\frac{0.02\sqrt{1 - \xi^2}}{\sqrt{1 - 2\xi\omega_n T_R + (\omega_n T_R)^2}}\right) \quad (25)$$

The key characteristic quantities derived from the analytical solution of the equivalent model form the physical information foundation for the subsequent physics–data hybrid prediction model. The values of these characteristic quantities are uniquely determined by the system’s equivalent parameter set, which essentially represents the comprehensive manifestation of the coordinated effects of operational conditions and all frequency regulation resource control systems. Therefore, the equivalent model developed in this study describes, through its identifiable equivalent parameters, the intrinsic properties of the system’s overall frequency response under given system structures and control strategies.

3. A Frequency Prediction Architecture Integrating WAMS Information with Physical Model

3.1. Principles of the MLP-GRU-Attention Model

To achieve accurate sequence prediction of power system transient frequency, a model capable of simultaneously processing high-dimensional spatial features, complex temporal dynamics, and maintaining physical interpretability must be constructed. For this purpose, this paper designs a hybrid MLP-GRU-Attention architecture, whose overall structure is illustrated in Figure 8. In this architecture, the MLP encoder is responsible for fusing and encoding the WAMS multi-bus measurement data and the key features from the physical model into a global feature vector. The GRU decoder is specifically designed to learn the complex temporal patterns arising from the dynamic responses of heterogeneous frequency regulation resources [32–34]. The attention mechanism dynamically focuses on the key features at different prediction stages during the decoding process. These three components work synergistically, forming a complete technical framework that progresses from “feature fusion” to “temporal learning” and finally to “dynamic focusing”, collectively enhancing the prediction accuracy and reliability for the entire transient frequency response process.

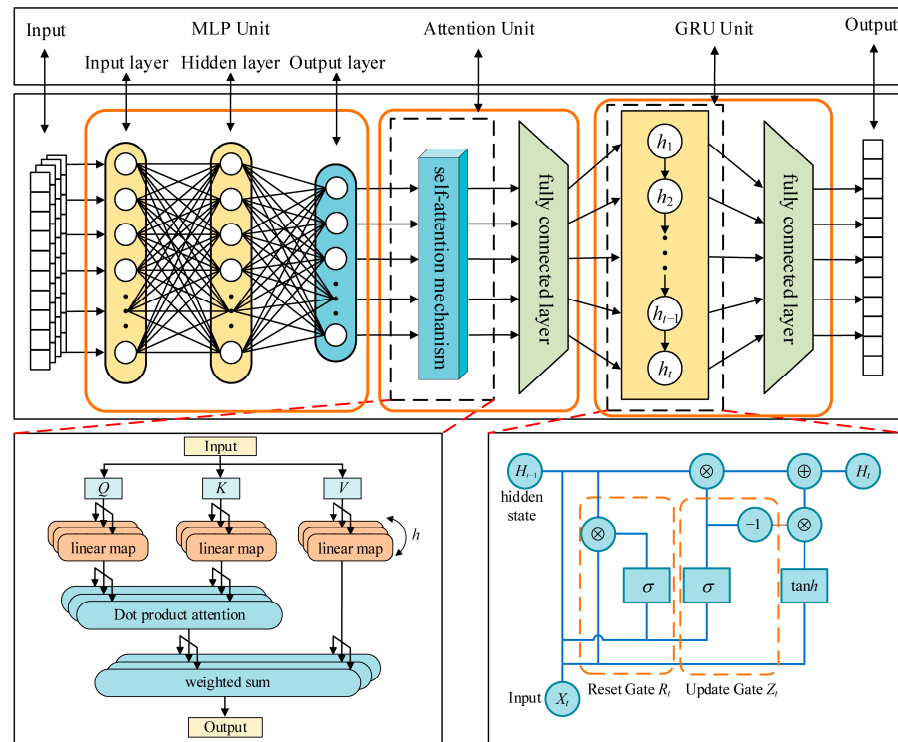


Figure 8. Schematic diagram of the MLP-GRU-Attention model structure.

(1) Encoder Design

The encoder employs a Multilayer Perceptron (MLP) as its core structure. Its input is the preprocessed and feature-weighted fused vector $X = [X_{WAMS}, X_{phy}]$, which consists of real-time grid snapshot data collected by WAMS—including node voltages (V), phase angles (θ), active power (P), and reactive power (Q)—and vector X_{phy} , which represents the physically critical features weighted by information entropy, such as the initial RoCoF Δf_{begin} , frequency nadir Δf_{min} , and steady-state frequency f_{ss} .

The MLP compresses and aggregates high-dimensional features through a three-layer nonlinear transformation, with the specific process described as follows:

(1) Input layer

The fused feature vector X is mapped into the hidden layer space, where its dimensionality is reduced from d_{in} to d_{h1} using the ReLU activation function, as follows:

$$H_1 = \text{ReLU}(W_1 X + b_1) \tag{26}$$

where $W_1 \in \mathbb{R}^{d_{h1} \times d_{in}}$ and $b_1 \in \mathbb{R}^{d_{h1}}$ denote learnable parameters.

(2) Hidden layer

The transformed features H_1 undergo further nonlinear transformation while maintaining the dimensionality d_{h1} , thereby enhancing the model’s capacity to capture complex nonlinear relationships. This process is expressed as follows:

$$H_2 = \text{ReLU}(W_2 H_1 + b_2) \tag{27}$$

where $W_2 \in \mathbb{R}^{d_{h1} \times d_{in}}$ and $b_2 \in \mathbb{R}^{d_{h1}}$ denote learnable parameters.

(3) Output layer

The input H_2 is encoded into a global feature vector X_{ctx} , with its dimensionality compressed to d_{ctx} . This vector comprehensively represents the initial state and correlations among key features following a grid disturbance, as expressed by the formula:

$$X_{ctx} = W_3 H_2 + b_3 \tag{28}$$

where $W_3 \in \mathbb{R}^{d_{h1} \times d_{in}}$ and $b_3 \in \mathbb{R}^{d_{h1}}$ denote learnable parameters.

(2) Decoder Design

The decoder employs a Gated Recurrent Unit (GRU) network to generate the future time-step frequency prediction sequence $\hat{y}_1, \hat{y}_2, \dots, \hat{y}_T$ in an autoregressive manner. By dynamically regulating the retention and forgetting of historical information through its update and reset gates, the GRU effectively mitigates the vanishing gradient problem during long-sequence training. The core computational procedure is as follows:

(1) Reset Gate r_t

It regulates the influence of the previous hidden state h_{t-1} on the current candidate hidden state h_t , as expressed by the formula:

$$r_t = \sigma(W_r[h_{t-1}, \hat{y}_{t-1}, X_{\text{ctx}}] + b_r) \quad (29)$$

where σ denotes the Sigmoid activation function, $h_{t-1} \in \mathbb{R}^{d_{h2}}$ represents the hidden state from the previous time step, \hat{y}_{t-1} corresponds to the predicted frequency value at the previous time step, and $W_r \in \mathbb{R}^{d_{h2} \times (d_{h2} + 1 + d_{\text{ctx}})}$ and $b_r \in \mathbb{R}^{d_{h2}}$ are learnable parameters.

(2) Update Door z_t

It controls the integration ratio between the historical hidden state h_{t-1} and the current candidate hidden state h_t , as expressed by the formula:

$$z_t = \sigma(W_z[h_{t-1}, \hat{y}_{t-1}, X_{\text{ctx}}] + b_z) \quad (30)$$

where $W_z \in \mathbb{R}^{d_{h2} \times (d_{h2} + 1 + d_{\text{ctx}})}$ and $b_z \in \mathbb{R}^{d_{h2}}$ denote learnable parameters.

(3) Candidate Hidden Status \tilde{h}_t

The hidden state is updated based on the historical information filtered by the reset gate and the current input, as follows:

$$\begin{cases} h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \\ \hat{y}_t = W_y h_t + b_y \end{cases} \quad (31)$$

where \odot denotes the element-wise multiplication operation, $W_{\tilde{h}} \in \mathbb{R}^{d_{h2} \times (d_{h2} + 1 + d_{\text{ctx}})}$ and $b_{\tilde{h}} \in \mathbb{R}^{d_{h2}}$ represent learnable parameters.

(4) Current hidden state h_t and predicted output \hat{y}_t

By combining the update gate weights, the current hidden state is generated and subsequently mapped to the frequency prediction value through a linear layer, as expressed by the formula:

$$\begin{cases} h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \\ \hat{y}_t = W_y h_t + b_y \end{cases} \quad (32)$$

where $W_y \in \mathbb{R}^{d_{h2} \times 1}$ and $b_y \in \mathbb{R}$ denote learnable parameters.

(3) Design of the Attention Mechanism

The attention mechanism serves as a critical component bridging the encoder and the decoder [35,36]. Its core function is to enable the decoder to dynamically focus on the features most relevant to the current prediction task within the global feature vector produced by the encoder when generating frequency predictions at each time step, thereby enhancing the model's sensitivity to key features and improving predictive interpretability. The specific computational steps are as follows:

(1) Computation of Attention Scores

The attention score e_t is computed based on the similarity between the hidden state h_{t-1} from the previous decoder time step and the global feature vector X_{ctx} from the encoder, using the scaled dot-product method. The formula is given as follows:

$$e_t = \frac{h_{t-1}^T X_{\text{ctx}}}{\sqrt{d_{\text{ctx}}}} \quad (33)$$

where $\sqrt{d_{\text{ctx}}}$ is a scaling factor introduced to mitigate the gradient saturation issue in the Softmax function caused by excessively large values after the dot product of high-dimensional vectors.

(2) Normalization of Attention Weights

The attention scores are converted into normalized weights α_t through the Softmax function, ensuring the sum of the weights equals 1, as follows:

$$\alpha_t = \frac{\exp(e_t)}{\sum_{k=1}^K \exp(e_k)} \quad (34)$$

where K denotes the number of segments of the global feature vector X_{ctx} , which are partitioned according to physical feature categories such as Δf_{begin} , Δf_{min} , and f_{ss} .

(3) Dynamic Global Feature Vector Generation

Based on the attention weights α_t , a weighted sum of the global feature vector X_{ctx} is computed, yielding the dynamic global feature vector x_t for the current time step, as follows:

$$x_t = \sum_{k=1}^K \alpha_t^k X_{\text{ctx}}^k \quad (35)$$

where α_t denotes the k -th segment of the global feature vector X_{ctx} , and x_t represents the corresponding attention weight for that segment.

3.2. Information Entropy-Based Weighting Method for Physical Features

The key features output by physical models exhibit varying information values across different grid disturbance scenarios. For instance, feature Δf_{min} demonstrates significant fluctuations during fault scenarios, containing rich transient stability information, whereas feature f_{ss} changes gradually in steady-state conditions, contributing less informational value. To quantify the information value of each physical feature and achieve adaptive weighting, this paper introduces information entropy theory, which characterizes the information content of a feature by its degree of uncertainty.

(1) Physical Feature Extraction

Based on the physics-based frequency response analysis model incorporating multi-resource frequency regulation developed in Section 1, the section data collected by WAMS is processed to compute and output N physical key features, forming the feature set $K_N = \{K_1, K_2, \dots, K_n\}$. To calculate the information entropy, this paper adopts the equal-frequency discretization method, dividing the value range of each feature K_i into m intervals such that each interval contains approximately the same number of historical samples, thereby reducing discretization error. Let the discrete intervals of feature K_i be denoted as $[x_{i1}, x_{i2}), [x_{i3}, x_{i4}), \dots, [x_{im}, x_{i(m+1)}]$, where x_{i1} and $x_{i(m+1)}$ represent the minimum and maximum values of the feature, respectively.

(2) Computation of Information Entropy and Weight Normalization

Information Entropy $H(K_i)$ is used to quantify the uncertainty of feature K_i across different scenarios. A higher entropy value indicates a more dispersed distribution of the feature's values, which corresponds to richer information content and a greater contribution to frequency prediction. Based on historical datasets, the probability p_{ij} of feature K_i falling into the j -th discrete interval (where $j = 1, 2, \dots, m$) is statistically derived. The information entropy $H(K_i)$ is calculated as follows:

$$H(K_i) = -\sum_{j=1}^m p_{ij} \log_2 p_{ij} \quad (36)$$

To address the issue of magnitude disparities in raw entropy values, a normalization procedure is applied to the information entropy of each feature, resulting in the feature weight ω_i and ensuring that the sum of all weights equals 1. The formula is defined as follows:

$$\omega_i = \frac{H(K_i)}{\sum_{i=1}^N H(K_i)} \quad (37)$$

(3) Weighted Physical Feature Vector Formulation

Each physical feature K_i is multiplied by its corresponding weight ω_i to obtain the weighted physical feature $K_{\text{phy}}^i = K_i \cdot \omega_i$. These weighted features are then integrated to form the weighted physical feature vector $X_{\text{phy}} = [K_{\text{phy}}^1, K_{\text{phy}}^2, \dots, K_{\text{phy}}^N]$. This vector is concatenated with the raw WAMS snapshot data vector X_{WAMS} , forming the input vector $X = [X_{\text{WAMS}}, X_{\text{phy}}]$ for the MLP encoder. This process achieves a deep fusion of system-level physical priors with spatiotemporal multi-bus measurement data, establishing an informational foundation for subsequent data-driven models that incorporates both physical consistency and spatial detail.

3.3. Construction of the Weighted Physical Feature Vector

To further strengthen the integration of physical knowledge with data-driven models and prevent predictions that violate grid physical laws, including sudden frequency jumps or steady-state frequency deviations beyond acceptable ranges, this paper proposes a physics-guided fusion architecture. This framework incorporates a physics-informed hybrid loss function that embeds power system operational constraints into the training process, ensuring both high accuracy in fitting historical data and strict adherence to physical requirements.

Conventional data-driven models solely employ the Mean Squared Error (MSE) to quantify the deviation between predicted and true values, which can lead to predictions that violate physical laws. To address this issue, this paper designs a hybrid loss function L , comprising a data fitting loss term L_{MSE} and a physical consistency loss term L_{phy} , formulated as follows:

$$L = \alpha \cdot L_{\text{MSE}} + \beta \cdot L_{\text{phy}} \quad (38)$$

where the hyperparameters α and β are adjustable.

(1) Data Fitting Loss L_{MSE}

This component measures the overall deviation between the model-predicted frequency sequence $\hat{Y} = [\hat{y}_1, \hat{y}_2, \dots, \hat{y}_T]$ and the true frequency sequence $Y = [y_1, y_2, \dots, y_T]$, ensuring the model's baseline prediction accuracy. It is defined as follows:

$$L_{\text{MSE}} = \frac{1}{T} \sum_{t=1}^T (\hat{y}_t - y_t)^2 \quad (39)$$

(2) Physics Consistency Loss

The design of L_{phy} is grounded in the core physical principles governing power system transient frequency response, targeting the following three key metrics to formulate constraint terms:

(1) Initial RoCoF Consistency Constraint

This constraint ensures that the deviation between the predicted initial RoCoF $\Delta \hat{f}_{\text{begin}} = (\hat{y}_2 - \hat{y}_1) / \Delta t$ and the value $\Delta f_{\text{begin}}^{\text{phy}}$ calculated by the physical model remains below a threshold $\epsilon_1 = 0.1$ Hz/s. It is formulated as follows:

$$L_{\Delta f_{\text{begin}}} = \max\left(0, \left| \Delta \hat{f}_{\text{begin}} - \Delta f_{\text{begin}}^{\text{phy}} \right| - \epsilon_1\right) \quad (40)$$

(2) Frequency Nadir Consistency Constraint

This constraint ensures that the deviation between the predicted value $\Delta \hat{f}_{\min} = \min(\hat{y}_1, \hat{y}_2, \dots, \hat{y}_T)$ and the value $\Delta f_{\min}^{\text{phy}}$ calculated by the physical model remains below a threshold $\epsilon_2 = 0.05$ Hz, formulated as follows:

$$L_{\Delta f_{\min}} = \max\left(0, \left| \hat{f}_{\Delta \min} - f_{\Delta \min}^{\text{phy}} \right| - \epsilon_2\right) \tag{41}$$

(3) Steady-State Trend Consistency Constraint

This constraint ensures that the deviation between the predicted steady-state frequency $\hat{f}_{\text{ss}} = \hat{y}_T$ (assuming steady state is reached at time T) and the value $f_{\text{ss}}^{\text{phy}}$ calculated by the physical model remains below the threshold $\epsilon_3 = 0.02$ Hz, formulated as follows:

$$L_{f_{\text{ss}}} = \max\left(0, \left| \hat{f}_{\text{ss}} - f_{\text{ss}}^{\text{phy}} \right| - \epsilon_3\right) \tag{42}$$

Thus, the comprehensive formula for the physics consistency loss L_{phy} is given by

$$L_{\text{phy}} = L_{\Delta \hat{f}_{\text{begin}}} + L_{\Delta f_{\min}} + L_{f_{\text{ss}}} \tag{43}$$

Through the aforementioned hybrid loss function, the model is required not only to fit historical data during training but also to satisfy physical constraints, thereby significantly enhancing the reliability and generalization capability of predictions. This approach is particularly suitable for scenarios where grid inertia fluctuates substantially due to high penetration of renewable energy integration.

4. Online Prediction Method for Transient Frequency Response Based on the Fusion of WAMS Data and Physical Model

Based on the frequency prediction architecture integrating WAMS information and physical model introduced in the previous section, this paper proposes an online prediction method for transient frequency response that fuses WAMS data with the physical model. The implementation process primarily consists of two key stages: offline training and online prediction. The specific workflow is illustrated in Figure 9.

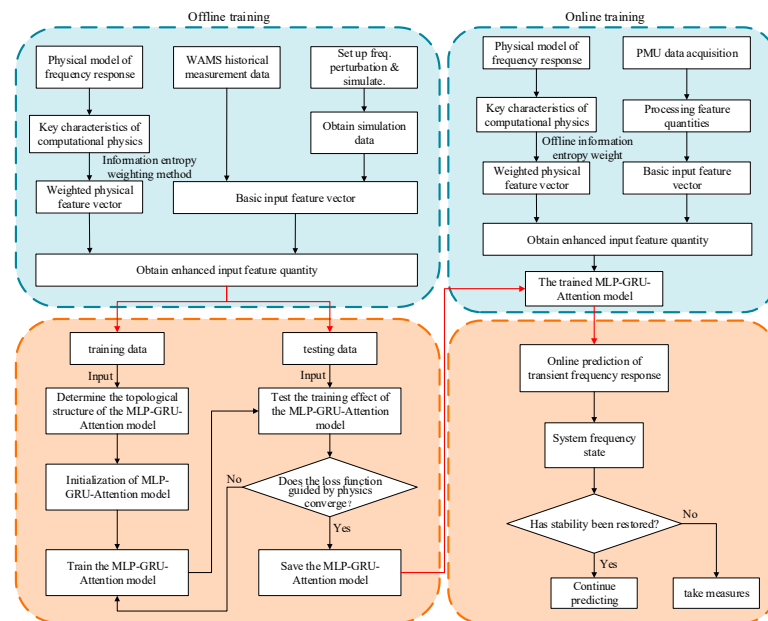


Figure 9. Flowchart of the online prediction method for transient frequency response based on WAMS data and physical model fusion.

(1) Offline Training

(1) Use historical measurement data from the renewable energy grid WAMS and time-domain simulation data from the frequency response physical model, incorporating multiple heterogeneous frequency regulation resources as raw sample data. Perform data cleaning, missing data repair, time alignment, and resampling on the raw samples to ensure temporal consistency and data integrity.

(2) Extract pre-disturbance steady-state grid cross-section data, instantaneous disturbance data, and post-disturbance initial-time data from the preprocessed WAMS historical measurement data and time-domain simulation data to form the basic input feature set.

(3) Based on the time-domain analytical solution for renewable grid frequency response and related formulas derived in Section 2.2, compute the physical key features for each sample. Combine these with the information entropy weighting method from Section 3.2 to obtain weighted physical feature vectors. Concatenate the basic input feature set with the weighted physical feature vectors to form the complete input set.

(4) Map the values of the complete input set to the interval $[0, 1]$ for normalization. This eliminates interference from dimensional differences among features during model training, ensuring convergence efficiency.

(5) Randomly partition the normalized input set into training and testing sets at a 2:1 ratio. Configure the parameters of the MLP-GRU-Attention model.

(6) After initializing the MLP-GRU-Attention model parameters, feed the training set into the model and train using the hybrid loss function designed in Section 3.3. Continue training until the model's loss value converges on the training set. If the prediction error on the test set does not meet the preset requirements, adjust the model hyperparameters. Iterate through the parameter tuning and model training process until the prediction error on the test set stabilizes within an acceptable range. This ultimately yields the fully trained physics-guided MLP-GRU-Attention frequency prediction model.

(2) Online Prediction

(1) WAMS data is collected in real time via phasor measurement units (PMU) extensively deployed across the grid and uploaded to the master station for rapid preprocessing.

(2) Based on real-time grid section characteristics, invoke the physical model equations from the offline training phase to rapidly compute real-time physical key features. Apply the offline-trained information entropy weights to these features, yielding a real-time weighted physical feature vector. Concatenate the real-time grid section characteristics with the weighted physical feature vector, normalize the combined input, and feed it into the offline-trained MLP-GRU-Attention frequency prediction model.

(3) The model outputs real-time transient frequency response prediction curves and key indicators. Based on the prediction results, it determines whether the system frequency is stable. If the frequency prediction indicates that the system can recover to stability, it continues collecting real-time data for subsequent frequency predictions. Otherwise, it immediately triggers emergency grid control measures to prevent system frequency instability.

5. Analysis of the Algorithm

5.1. Example System

To validate the effectiveness of the proposed method, this paper establishes a detailed time-domain simulation model in PSS/E that emulates a realistic renewable energy-integrated power grid. The pre- and post-disturbance data obtained from the time-domain simulations are utilized as the WAMS data required by the proposed approach. For this purpose, the conventional IEEE 39-bus system is modified by incorporating equivalent wind farms at buses 30, 37, and 39, equivalent photovoltaic power stations at buses 8, 33, and 34,

and DC links at buses 16 and 26. This transforms the grid into a typical power system with high penetration of renewable energy, whose structure is illustrated in Figure 10.

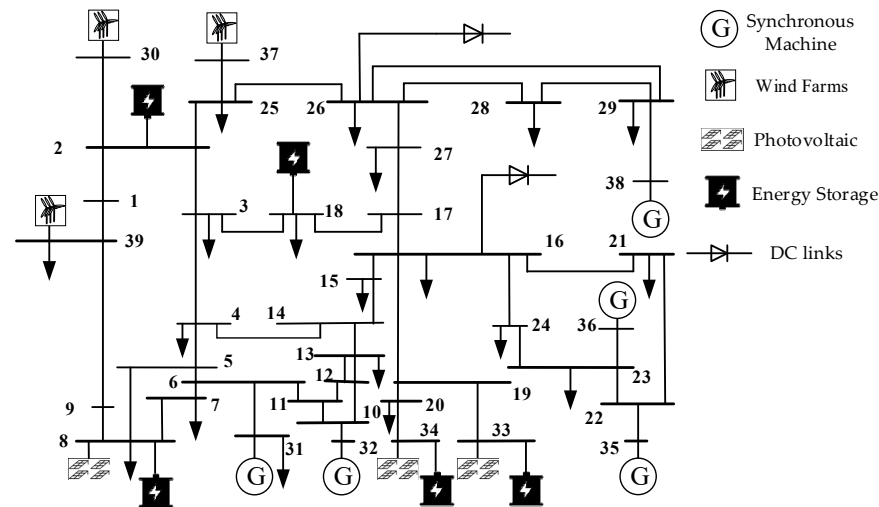


Figure 10. Improved IEEE39 node grid structure diagram.

The MLP-GRU-Attention model hyperparameters are set as follows: 2-layer MLP encoder and 1-layer GRU decoder structure, 4-head attention mechanism, hidden layer dimension 64, feed-forward network dimension 256 (ReLU activation function is used), batch size 32, initial learning rate 1×10^{-4} (Adam's optimizer), Dropout rate 0.2, and number of training rounds, i.e., 150 rounds (with early stopping method). The input features are Min–Max normalized to the interval [0, 1].

In this paper, Python 3.9 programming is used to invoke the PSS/E simulation platform to carry out cyclic simulation analysis of the aforementioned new energy grid, where load increase perturbation is carried out for a single load node or multiple load nodes. The size of the load perturbation is randomly set in the range of 0%~50%, and the operation data of each node before and immediately after load perturbation is proposed to be the basic data of the WAMS, so as to obtain 9000 groups of samples, including 6000 groups of training samples and 3000 groups of testing samples.

5.2. Performance Evaluation Index

To comprehensively evaluate the performance of the proposed prediction model from multiple dimensions, this paper selects four evaluation metrics with complementary physical significance: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Coefficient of Determination (R^2), and Dynamic Time Warping (DTW). These metrics provide quantitative assessments from the perspectives of absolute error level, sensitivity to large errors, goodness of trend fit, and transient process shape similarity, respectively. Their definitions are as follows:

(1) Mean Absolute Error (MAE)

MAE calculates the average of absolute differences between predicted values and true values, which can intuitively reflect the overall level of prediction errors.

$$V_{\text{MAE}} = \frac{1}{N} \sum_{i=1}^N |y_i - f(x_i)| \quad (44)$$

(2) Root Mean Square Error (RMSE)

RMSE is the square root of the average of squared errors, which is more sensitive to larger errors in predictions and can effectively measure prediction accuracy and stability.

$$V_{RMSE} = \left[\frac{1}{N} \sum_{i=1}^N (y_i - f(x_i))^2 \right]^{1/2} \tag{45}$$

(3) Coefficient of Determination (R^2)

R^2 measures the proportion of variance in the observed data that is explained by the predicted results. A value closer to 1 indicates a better goodness-of-fit of the model to the system dynamics.

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - f(x_i))^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \tag{46}$$

(4) Dynamic Time Warping (DTW)

DTW calculates the minimum cumulative distance by finding the optimal nonlinear alignment between two time series, making it particularly suitable for evaluating the shape similarity of transient process curves that may have phase differences in the time axis.

$$DTW(X, Y) = D(n, m) \tag{47}$$

where N is the total number of time steps; y_i represents the actual value at the i -th time step; $f(x_i)$ denotes the predicted value at the i -th time step, where $DTW(X, Y)$ represents the Dynamic Time Warping distance between sequence X and sequence Y ; D is the cumulative distance matrix; the length of X is n , and the length of Y is m .

5.3. Prediction Performance Analysis

In order to verify the effectiveness of the method proposed in this paper, for the method in this paper, the pure physical model prediction method, the pure MLP-GRU-Attention model prediction method, the existing DBN model prediction method and the PSS/E time-domain simulation method, a sample is randomly selected among the test samples to carry out a comparative analysis of prediction effects under different methods, and the corresponding prediction results of its transient frequency are as follows The corresponding prediction results of transient frequency are shown in Figure 11.

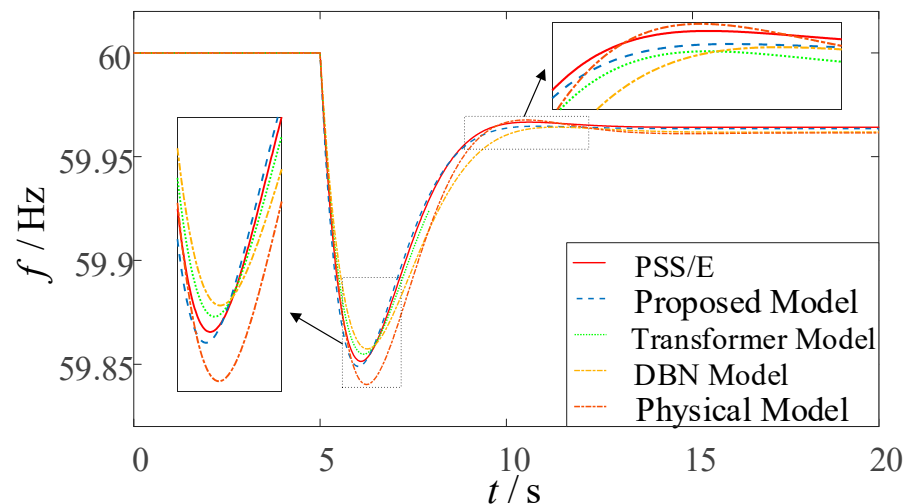


Figure 11. Comparison of transient frequency response curves under different prediction methods.

As illustrated in Figure 11, the transient frequency response curve predicted by the proposed method demonstrates the closest agreement with the benchmark PSS/E simulation curve, accurately reproducing the complete dynamic process of frequency decline and recovery. During the initial stage following the disturbance, the transient frequency response curves predicted by all methods remain relatively similar. However, as time progresses, the frequency curves exhibit significant variations, and the discrepancies between the predictions of different methods become more pronounced. To evaluate the predictive performance of each method regarding key frequency response characteristics, a comparative statistical analysis of these key parameters was conducted, and the Absolute Errors (AE) were calculated, as summarized in Table 2 below.

Table 2. Comparison of key transient frequency characteristics under different prediction methods.

Methodologies	Initial Frequency Rate of Change (Hz/s)		Minimum Frequency (Hz)		Minimum Frequency Arrival Time (s)		Steady State Frequency (Hz)	
	Δf_{begin}	AE	Δf_{min}	AE	t_z	AE	f_{ss}	AE
Methodology of this paper	0.124	0.004	59.849	0.003	6.02	0.07	59.963	0.001
MLP-GRU-Attention modeling approach	0.111	0.009	59.859	0.007	6.17	0.08	59.959	0.005
DBN modeling approach	0.116	0.004	58.864	0.012	6.30	0.21	59.962	0.002
physical model	0.131	0.011	59.836	0.016	6.31	0.22	59.961	0.003

As shown in Table 2, regarding the RoCoF, frequency nadir, time to reach frequency nadir, and steady-state frequency, the prediction accuracy of the proposed method for RoCoF shows improvements of 55.56%, 0%, and 63.6% compared to the purely MLP-GRU-Attention-based model, the purely DBN-based method, and the purely physics-based method, respectively. The prediction accuracy for frequency nadir is improved by 57.14%, 75%, and 98.9%; for the time to reach frequency nadir by 12.50%, 66.67%, and 68.18%; and for steady-state frequency by 80.00%, 50.00%, and 66.67%, respectively. These results demonstrate that the proposed method not only exhibits overall superior prediction performance compared to conventional approaches but also achieves higher accuracy in predicting key frequency characteristics. By integrating frequency response features with precisely measured data, the method achieves outstanding prediction precision. In particular, the RoCoF derived from the frequency response ensemble model effectively guides the model in capturing early-stage frequency dynamics, while the incorporated physical constraints limit the deviation between the predicted nadir and steady-state frequency within 0.003 Hz, further ensuring the reliability of the prediction results.

To thoroughly validate the generalizability of the conclusions drawn from the aforementioned case analysis, the proposed fusion model, the standalone MLP-GRU-Attention model, and the DBN model were trained using the training dataset. Subsequently, 100 test samples were randomly selected from the original test set to evaluate the performance of each model in predicting post-disturbance transient frequency, with the computational results from the physical model also included for comparison. All predictive performance assessments were benchmarked against the PSS/E time-domain simulation results described in Section 5.1. Based on this benchmark, the average values of four metrics—RMSE, MAE, R^2 , and DTW—across all samples were calculated. The comprehensive predictive performance comparison of the four methods is ultimately presented in Figure 12.

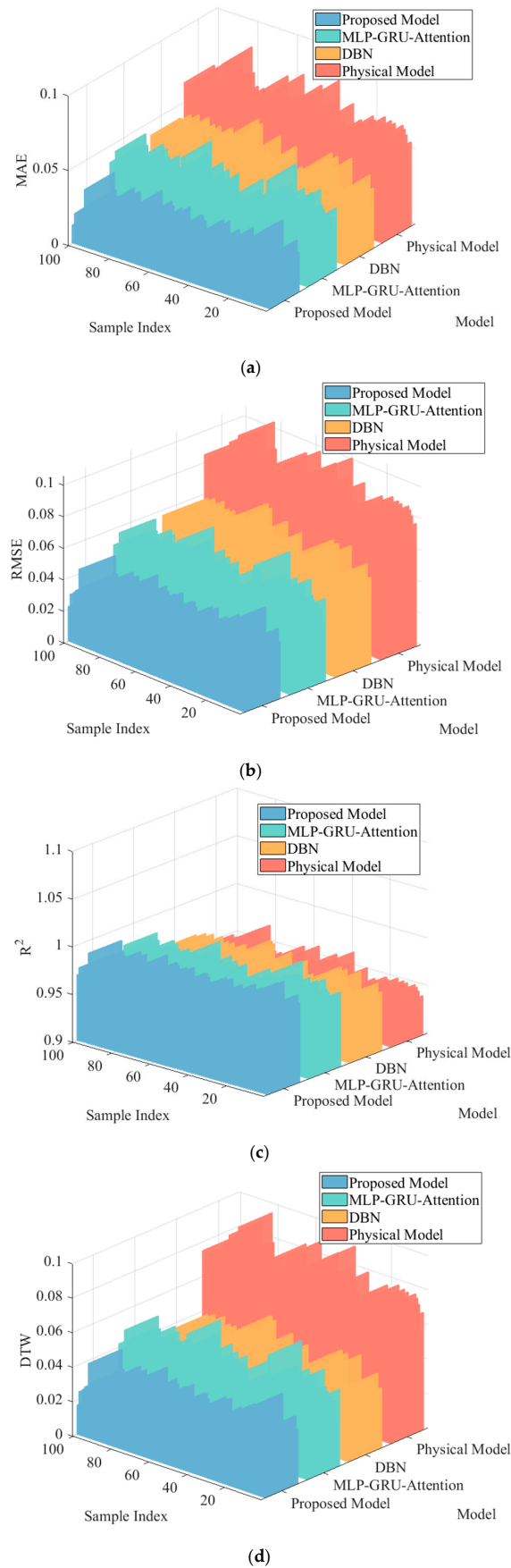
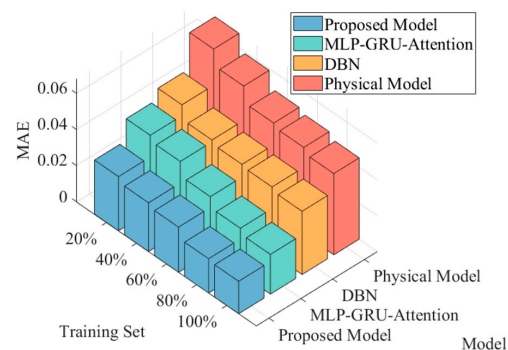


Figure 12. Comparison of model prediction performance of different methods on a random test sample set. **(a)** Comparison of MAE indicators. **(b)** Comparison of RMSE indicators. **(c)** Comparison of R² indicators. **(d)** Comparison of DTW indicators.

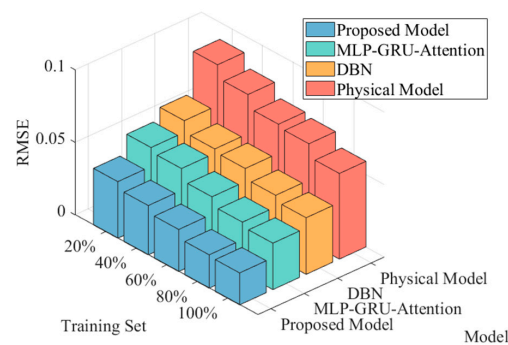
As observed in Figure 12, the proposed method consistently demonstrates superior comprehensive performance across all four core evaluation metrics when benchmarked against PSS/E time-domain simulation results. The simultaneous significant reductions in both MAE and RMSE substantiate enhanced prediction accuracy from two perspectives: measuring average absolute deviation and penalizing large errors. Particularly, the minimized DTW distance indicates that the frequency curve predicted by our method exhibits the closest dynamic morphology to the actual curve, accurately capturing the temporal characteristics of the transient process. Although the absolute improvement in R^2 appears modest, its value being closest to unity reflects the model's optimal goodness-of-fit in representing the overall variation trend of frequency response, revealing its exceptional capability in capturing underlying physical mechanisms. These metrics collectively validate the comprehensive superiority of our fusion approach through multiple complementary dimensions, including error magnitude, curve morphology, and trend fitting.

5.4. Evaluation of Model Generalization Ability for Small Samples

In order to verify the generalization ability of the data-model fusion method proposed in this paper, 20%, 40%, 60%, 80%, and 100% of the total training samples are randomly selected as five new training sample sets, and the models of this paper's method, MLP-GRU-Attention model prediction method, DBN model prediction method, and the physical model method are constructed for the comparative analysis of the prediction effect under the five training sets. The final comparison of the prediction effect of the four methods is shown in Figure 13.

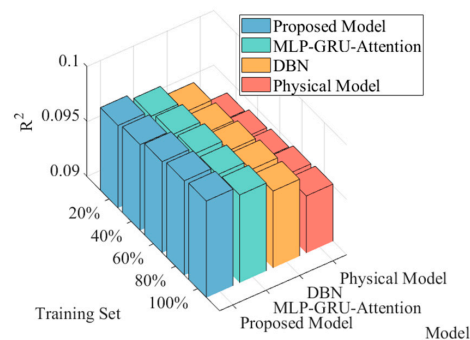


(a)

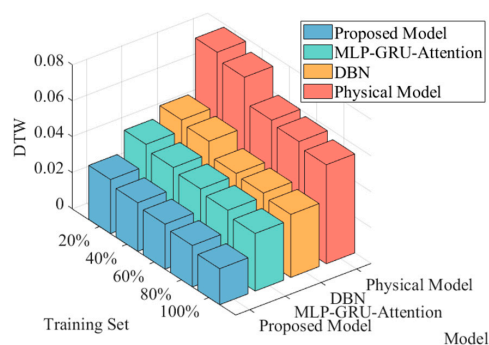


(b)

Figure 13. Cont.



(c)



(d)

Figure 13. Comparison of model prediction performance of different methods under different training sets. (a) Comparison of MAE indicators. (b) Comparison of RMSE indicators. (c) Comparison of R^2 indicators. (d) Comparison of DTW indicators.

As observed in Figure 13, with the increase in training sample size, the prediction error metrics of all models show a decreasing trend. Even under the limited 20% training sample condition, the proposed method maintains relatively lower values in MAE, RMSE, and DTW metrics with the smallest performance degradation, highlighting the effective enhancement of model generalization capability in small-sample scenarios through physics-informed knowledge embedding. This phenomenon can be attributed to the physical guidance mechanism providing prior knowledge to the model, reducing its dependency on pure data volume, thereby enabling effective learning of complex nonlinear mapping relationships even with scarce data. In summary, the proposed model demonstrates superior prediction accuracy and data utilization efficiency across all four core evaluation metrics, with these advantages becoming particularly prominent under insufficient training sample conditions.

5.5. Model Noise Resistance Evaluation

Considering the actual engineering applications, the measurement data set of WAMS inevitably contains Gaussian white noise, impulse interference, and other composite error sources due to the limitation of sensor accuracy and the influence of communication channel interference in the data acquisition process. In order to evaluate the noise immunity of the model, this paper adds 5%, 10%, 15%, 20%, 25% and 30% Gaussian noise signals to the original test sample time series data to simulate the actual situation of WAMS measurement data. In order to verify the effectiveness of this paper's method, it is compared with the purely physical model prediction method, the pure MLP-GRU-Attention model prediction method, and the existing DBN model prediction method under different noise levels, and the results of the evaluation index comparison are shown in Figure 14 below.

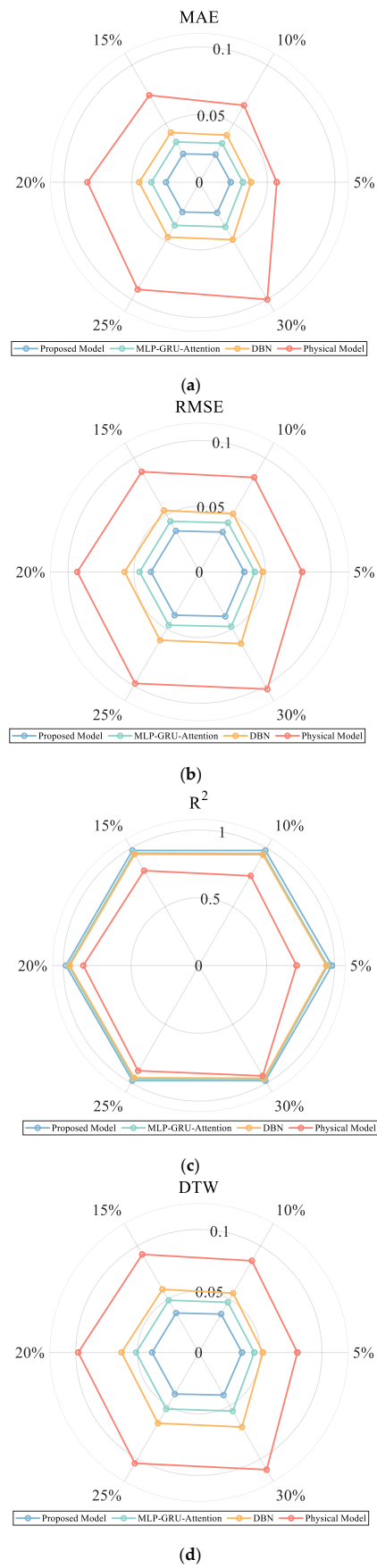


Figure 14. Comparison of model prediction performance of different methods under different Gaussian noise effects. **(a)** Comparison of MAE indicators. **(b)** Comparison of RMSE indicators. **(c)** Comparison of R^2 indicators. **(d)** Comparison of DTW indicators.

As shown in Figure 14, the proposed method demonstrates the most gradual increase in all evaluation metrics under progressively increasing noise levels. The slow growth of both MAE and RMSE quantitatively confirms the method's insensitivity to measurement noise in terms of absolute error and prediction volatility, indicating excellent robustness. In contrast, traditional physics-based methods exhibit significantly rising evaluation metrics with increasing noise, primarily because their equivalent parameters are directly calculated from noisy data without establishing noise compensation mechanisms. The strong noise resistance of our method can be attributed to two key aspects: first, the data-driven module inherently possesses certain noise-smoothing capabilities; second, and more importantly, the embedded physics-guided framework enhances the contribution of key physical features through information entropy weighting and constrains prediction trajectories via physical consistency loss. This essentially enables adaptive noise filtering and trajectory calibration, thereby significantly improving reliability in challenging measurement environments.

6. Conclusions

Based on WAMS technology, this paper proposes a transient frequency response online prediction method that deeply integrates the WAMS data-driven model with the physical model. By constructing a physical model of frequency response that accounts for multiple types of FM resources, key features with clear physical significance are extracted; information entropy theory is introduced to realize adaptive weighted fusion of physical features; a physically guided MLP-GRU-Attention prediction framework is designed, and physical consistency constraints are introduced into the loss function to ensure that the prediction results are strictly in line with the dynamic laws of the system. Simulation validation based on the improved IEEE39-node system shows that the proposed method significantly outperforms the traditional data-driven and purely physical model methods in terms of prediction accuracy, generalization ability, and anti-noise performance, and the conclusions are as follows:

- (1) The physics–data fusion method proposed in this paper effectively integrates the mechanistic interpretability of physical models with the high-precision learning capability of data-driven approaches through information entropy-weighted fusion of system-level key features extracted from the average frequency model and multi-bus measurement data provided by WAMS. As evidenced by simulation results, the proposed method demonstrates highly accurate reproduction of actual system frequency dynamics benchmarked against PSS/E simulations, exhibiting superior accuracy and reliability throughout the entire transient frequency response prediction process.
- (2) By introducing the information entropy-weighted physical feature fusion and the physics-guided machine learning framework, the generalization ability of the model under small sample conditions is significantly improved, and its applicability in extreme disturbance scenarios is enhanced.
- (3) The proposed method shows good robustness in the presence of WAMS measurement noise, and the embedding of physical knowledge effectively suppresses noise interference, maintains the stability and consistency of the prediction results, and provides reliable technical support for the online sensing and control of frequency security in new energy power grids.

Author Contributions: Y.H. spearheaded the research conceptualization and the design of the online transient frequency prediction framework that deeply integrates WAMS data with a physical model, led the development of the physics-guided machine learning algorithm (including the MLP-GRU-Attention model), and drafted the initial manuscript. K.Y. constructed the multi-resource integrated frequency response physical model (covering conventional thermal units, wind power, PV, energy storage, and HVDC transmission), derived the time-domain analytical solution of transient

frequency and the calculation method of key features (e.g., RoCoF, frequency nadir), and critically revised the technical content of the manuscript. H.X. conducted simulation validation based on the modified IEEE 39-bus system, analyzed the prediction performance metrics (such as MAE, RMSE, R^2 , and DTW) and the generalization ability under small-sample conditions, and contributed to the interpretation of experimental results. T.H. provided cross-institutional research support (relying on the Department of Energy, Politecnico di Torino), optimized the integration logic of physical model transfer functions and WAMS snapshot data, and participated in the technical review of the frequency regulation resource modeling. Y.L. and T.W. jointly curated historical WAMS measurement datasets and time-domain simulation samples, performed parameter tuning for the MLP-GRU-Attention model (e.g., hidden layer dimension, learning rate), and verified the model's noise immunity under different Gaussian noise intensities. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Sichuan Major Science and Technology Special Project: Key Technologies and Demonstration Application of Friendly Integration and Intelligent Regulation for Clean Energy, grant number 2025ZDZX0033.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data supporting the findings of this study are subject to third-party restrictions and are not publicly available at this time. Upon reasonable request and in compliance with confidentiality agreements, data access inquiries should be directed to the corresponding author, Yi Hu.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

FM	Frequency Modulation
SFR	System Frequency Response
WAMS	Wide-Area Measurement System
SME	Single-Machine Equivalent
HVDC	High Voltage Direct Current Transmission
RoCoF	Rate of Change of Frequency
UFLS	Under-Frequency Load Shedding
FLC	Frequency Limiting Controller
PI	Proportional-Integral
PV	Photovoltaic
MLP	Multilayer Perceptron
GRU	Gated Recurrent Unit
DBN	Deep Belief Network
RNN	Recurrent Neural Network
1D-CNN	One-Dimensional Convolutional Neural Network
TFAM	Temporal-Feature Attention Module
FSPM	Frequency Security Predictor Model
FDL	Frequency Danger Level
TSM	Time Security Margin
LSTM	Long Short-Term Memory
WOA	Whale Optimization Algorithm
NAR	Nonlinear AutoRegressive
MSE	Mean Squared Error
MAE	Mean Absolute Error
RMSE	Root Mean Square Error
MRE	Mean Relative Error

R ²	Coefficient of Determination
DTW	Dynamic Time Warping
PMU	Phasor Measurement Unit

References

1. Yu, L.; Zhang, L.; Meng, G.; Zhang, F.; Liu, W. Research on Multi-Objective Reactive Power Optimization of Power Grid with High Proportion of New Energy. *IEEE Access* **2022**, *10*, 116443–116452. [[CrossRef](#)]
2. Luo, Y.; Cai, L.; Zhang, N. Doubly-Fed Pumped Storage Units Participation in Frequency Regulation Control Strategy for New Energy Power Systems Based on Model Predictive Control. *Energy Eng.* **2025**, *122*, 765–783. [[CrossRef](#)]
3. Dhara, P.K.; Rather, Z.H. A Novel Approach to Determine and Maintain Area-Wise Minimum Inertia in Renewable Energy Dominated Power Systems. *IEEE Trans. Sustain. Energy* **2025**, *16*, 1084–1097. [[CrossRef](#)]
4. Cao, Y.; Wu, Q.; Zhang, H.; Li, C.; Zhang, X. Chance-Constrained Optimal Configuration of BESS Considering Uncertain Power Fluctuation and Frequency Deviation Under Contingency. *IEEE Trans. Sustain. Energy* **2022**, *13*, 2291–2303. [[CrossRef](#)]
5. Hao, L.; Ji, J.; Xie, D.; Wang, H.; Li, W.; Asaah, P. Scenario-based Unit Commitment Optimization for Power System with Large-scale Wind Power Participation in primary Frequency Regulation. *J. Mod. Power Syst. Clean Energy* **2022**, *8*, 1259–1267. [[CrossRef](#)]
6. Alves, E.F.; Polleux, L.; Guerassimoff, G.; Korpås, M.; Tedeschi, E. Allocation of Spinning Reserves in Autonomous Grids Considering Frequency Stability Constraints and Short-Term Solar Power Variations. *IEEE Access* **2023**, *11*, 29896–29908. [[CrossRef](#)]
7. Dai, D.; Liu, J. Human powered wireless charger for low-power mobile electronic devices. *IEEE Trans. Consum. Electron.* **2012**, *58*, 767–774. [[CrossRef](#)]
8. Tan, D.; Novosel, D. Energy challenge, power electronics & systems (PEAS) technology and grid modernization. *CPSS Trans. Power Electron. Appl.* **2017**, *2*, 3–11. [[CrossRef](#)]
9. Meng, L.; Zafar, J.; Khadem, S.K.; Collinson, A.; Murchie, K.C.; Coffele, F.; Burt, G.M. Fast Frequency Response From Energy Storage Systems—A Review of Grid Standards, Projects and Technical Issues. *IEEE Trans. Smart Grid* **2020**, *11*, 1566–1581. [[CrossRef](#)]
10. Zhang, Z.; Zhou, M.; Wu, Z.; Liu, S.; Guo, Z.; Li, G. A Frequency Security Constrained Scheduling Approach Considering Wind Farm Providing Frequency Support and Reserve. *IEEE Trans. Sustain. Energy* **2022**, *13*, 1086–1100. [[CrossRef](#)]
11. Azman, S.K.; Isbeih, Y.J.; Moursi, M.S.E. A Unified Online Deep Learning Prediction Model for Small Signal and Transient Stability. *IEEE Trans. Power Syst.* **2020**, *35*, 4585–4598. [[CrossRef](#)]
12. Zhang, J.; Wang, Y.; Zhou, G.; Wang, L.; Li, B.; Li, K. Integrating Physical and Data-Driven System Frequency Response Modelling for Wind-PV-Thermal Power Systems. *IEEE Trans.* **2024**, *39*, 217–228. [[CrossRef](#)]
13. Belforte, P.; Spina, D.; Lombardi, L.; Antonini, G.; Dhaene, T. Automated Framework for Time-Domain Piecewise-Linear Fitting Method Based on Digital Wave Processing of S-Parameters. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2020**, *67*, 235–248. [[CrossRef](#)]
14. Wu, F.; Qian, J.; Ju, P.; Zhang, X.; Jin, Y.; Xu, D.; Sterling, M. Transfer function based equivalent modeling method for wind farm. *J. Mod. Power Syst. Clean Energy* **2019**, *7*, 549–557. [[CrossRef](#)]
15. Wang, W.; Wang, Z.; Liu, X.; Li, W.; Li, Q.; Zhang, Y.; Chen, Q.; Guo, S.; Xu, Z. Frequency response mode prediction of power system after large disturbances based on deep belief neural network. *IEEE Access* **2023**, *11*, 113653–113666. [[CrossRef](#)]
16. Zhang, J.; Wang, Y.; Li, H.; Zhou, G.; Li, B.; Wang, L.; Li, K. SFR modeling for hybrid power systems based on deep transfer learning. *IEEE Trans. Ind. Inform.* **2023**, *20*, 399–410. [[CrossRef](#)]
17. Li, L.; Wu, J.; Zhao, P.; Li, B.; Shi, F.; Wang, Y. Online transient frequency safety prediction machine of power system based on time-feature attention module. *IEEE Trans. Power Syst.* **2022**, *38*, 3952–3964. [[CrossRef](#)]
18. Li, X.; Xiao, N.; Peng, B.; Ai, Z.; Wang, Y. Frequency prediction after disturbance of grid-connected wind power systems based on WOA and Attention-LSTM. *Energy Rep.* **2023**, *9*, 208–216. [[CrossRef](#)]
19. Ahakonye, L.A.C.; Nwakanma, C.I.; Lee, J.-M.; Kim, D.-S. Low computational cost convolutional neural network for smart grid frequency stability prediction. *Internet Things* **2024**, *25*, 101086. [[CrossRef](#)]
20. Chamorro, H.R.; Orjuela-Cañón, A.D.; Ganger, D.; Persson, M.; Gonzalez-Longatt, F.; Alvarado-Barrios, L.; Sood, V.K.; Martinez, W. Data-Driven Trajectory Prediction of Grid Power Frequency Based on Neural Models. *Electronics* **2021**, *10*, 151. [[CrossRef](#)]
21. Xu, L.; Li, L.; Wang, M.; Wang, X.; Li, Y.; Li, W.; Zhou, K. Online Prediction Method for Power System Frequency Response Analysis Based on Swarm Intelligence Fusion Model. *IEEE Access* **2023**, *11*, 13519–13532. [[CrossRef](#)]
22. Wang, Q.; Li, F.; Tang, Y.; Xu, Y. Integrating Model-Driven and Data-Driven Methods for Power System Frequency Stability Assessment and Control. *IEEE Trans. Power Syst.* **2019**, *34*, 4557–4568. [[CrossRef](#)]
23. Wang, S.; Wang, R.; Yang, J.; Hao, L. Online Incremental Dynamic Modeling Using Physics-Informed Long Short-Term Memory Networ for the Pneumatic Artificial Muscle. *IEEE Robot. Autom. Lett.* **2024**, *9*, 8435–8442. [[CrossRef](#)]
24. Tan, Y.; Chen, Y.; Li, Y.; Cao, Y. Linearizing Power Flow Model: A Hybrid Physical Model-Driven and Data-Driven Approach. *IEEE Trans. Power Syst.* **2020**, *35*, 2475–2478. [[CrossRef](#)]
25. Remon, D.; Cantarellas, A.M.; Rodriguez, P. Equivalent Model of Large-Scale Synchronous Photovoltaic Power Plants. *IEEE Trans. Ind. Appl.* **2016**, *52*, 5029–5040. [[CrossRef](#)]

26. Yang, L.; Zhou, N.; Zhou, G.; Chi, Y.; Chen, N.; Wang, L.; Wang, Q.; Chang, D. Day-ahead Optimal Dispatch Model for Coupled System Considering Ladder-type Ramping Rate and Flexible Spinning Reserve of Thermal Power Units. *J. Mod. Power Syst. Clean Energy* **2022**, *10*, 1482–1493. [[CrossRef](#)]
27. Sun, M.; Min, Y.; Chen, L. Optimal auxiliary frequency control of wind turbine generators and coordination with synchronous generators. *CSEE J. Power Energy Syst.* **2021**, *7*, 78–85.
28. Kakimoto, N.; Takayama, S.; Satoh, H. Power Modulation of Photovoltaic Generator for Frequency Control of Power System. *IEEE Trans. Energy Convers.* **2009**, *24*, 943–949. [[CrossRef](#)]
29. Hoke, A.F.; Shirazi, M.; Chakraborty, S. Rapid Active Power Control of Photovoltaic Systems for Grid Frequency Support. *IEEE J. Emerg. Sel. Top. Power Electron.* **2017**, *5*, 1154–1163. [[CrossRef](#)]
30. Muyeen, S.M.; Takahashi, R.; Tamura, J. Operation and Control of HVDC-Connected Offshore Wind Farm. *IEEE Trans. Sustain. Energy* **2010**, *1*, 30–37. [[CrossRef](#)]
31. Dai, J.; Tang, Y.; Wang, Q.; Jiang, P. Aggregation Frequency Response Modeling for Wind Power Plants with Primary Frequency Regulation Service. *IEEE Access* **2019**, *7*, 108561–108570. [[CrossRef](#)]
32. Zhang, B.; Xiong, D.; Xie, J.; Su, J. Neural Machine Translation with GRU-Gated Attention Model. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *31*, 4688–4698. [[CrossRef](#)]
33. Hasanat, S.M.; Ullahet, K.; Yousafal, H.; Munir, K.; Abid, S.; Bokhari, S.A.S.; Aziz, M.M.; Naqvi, S.F.M.; Ullah, Z. Enhancing Short-Term Load Forecasting with a CNN-GRU Hybrid Model: A Comparative Analysis. *IEEE Access* **2024**, *12*, 184132–184141. [[CrossRef](#)]
34. Chen, W.; Zhou, H.; Cheng, L.; Xia, M. Wind Turbine Blade Icing Diagnosis Using Convolutional LSTM-GRU with Improved African Vultures Optimization. *IEEE Open J. Instrum. Meas.* **2022**, *1*, 1–9. [[CrossRef](#)]
35. Yang, J.; He, H.; Zhao, X.; Wang, J.; Yao, T.; Cao, H.; Wan, M. Day-Ahead PV Power Forecasting Model Based on Fine-Grained Temporal Attention and Cloud-Coverage Spatial Attention. *IEEE Trans. Sustain. Energy* **2024**, *15*, 1062–1073. [[CrossRef](#)]
36. Bian, J.; Wang, L.; Scherer, R.; Wozniak, M.; Zhang, P.; Wei, W. Abnormal Detection of Electricity Consumption of User Based on Particle Swarm Optimization and Long Short Term Memory with the Attention Mechanism. *IEEE Access* **2021**, *34*, 1653–1656. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.