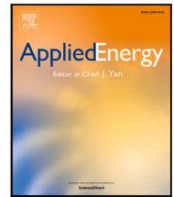


Listas de conteúdo disponíveis em [ScienceDirect](#)

Energia Aplicada

página inicial da revista: www.elsevier.com/locate/apenergy

Uma revisão da velocidade do vento e previsão de energia eólica com redes neurais profundas

Yun Wang ^a, Runmin Zou ^{a, *}, Fang Liu ^a, Lingjun Zhang ^b, Qianyi Liu ^a^a Escola de Automação, Universidade Centro-Sul, Changsha, Província de Hunan,^b Escola de Ciência da Computação, Universidade Hangzhou Dianzi, Hangzhou, Província de Zhejiang, China

INFORMAÇÕES DO ARTIGO

Palavras-chave:

Previsão da velocidade do vento

Previsão de energia eólica

Rede neural profunda

Pré-processamento de dados

Extração de recursos

Aprendizagem de relacionamento

RESUMO

O uso da energia eólica, uma forma de energia renovável e livre de poluição, para gerar eletricidade tem atraído cada vez mais atenção. No entanto, a geração intermitente de eletricidade resultante da natureza aleatória da velocidade do vento apresenta desafios para a segurança e estabilidade das redes de energia elétrica quando a energia eólica é integrada às redes em grande escala. Portanto, a previsão precisa da velocidade do vento e da potência eólica (WS/WP) gradualmente assumiu um papel fundamental na redução das flutuações da energia eólica no planejamento do despacho do sistema. Com o desenvolvimento de tecnologias de inteligência artificial, especialmente aprendizado profundo, um número crescente de modelos baseados em aprendizado profundo está sendo considerado para previsão de WS/WP devido à sua capacidade superior de lidar com problemas não lineares complexos. Este artigo analisa de forma abrangente as várias tecnologias de aprendizado profundo usadas na previsão de WS/WP, incluindo os estágios de processamento de dados, extração de recursos e aprendizado de relacionamento. O desempenho de previsão de alguns modelos populares é testado e comparado usando dois conjuntos de dados de vento do mundo real. Nesta revisão, são identificados três desafios para a previsão precisa de WS/WP sob condições complexas, a saber, incertezas de dados, recursos incompletos e relacionamentos não lineares intrincados. Além disso, as futuras direções de pesquisa são resumidas como um guia para melhorar a precisão das previsões de WS/WP.

1. Introdução

Com o rápido desenvolvimento econômico e o aumento contínuo dos padrões de vida, a demanda humana por energia também está aumentando [1]. No entanto, o uso de energia tradicional baseada em fósseis, como carvão, petróleo e gás natural, libera poluição que danifica o meio ambiente e causa o aquecimento global [1]. Além disso, devido à sua natureza não renovável e reservas limitadas, a exploração excessiva de combustíveis fósseis levará ao esgotamento dos recursos energéticos. Portanto, para resolver a crise energética e os problemas ambientais, é necessária energia renovável [2,3]. A energia eólica é uma forma de energia renovável, livre de poluição e amplamente distribuída, que recebeu ampla atenção global [4,5].

De acordo com o *Global Wind Report 2021* publicado pelo Global Wind Energy Council [6], cerca de 93 GW de novas instalações de energia eólica (WP) foram construídas em 2020 (conforme mostrado na Fig. 1(a)), um crescimento de 53% em relação a 2019. Isso elevou a capacidade total instalada de WP para 743 GW em 2020, um crescimento de 14,3% em relação ao ano anterior [6]. Com base nos dados das *Estatísticas de Capacidade Renovável 2021* fornecidos pela Agência Internacional de Energia Renovável [7], os 10 principais países por capacidade eólica acumulada em 2020 estão representados na Fig. 1(b), a partir da qual se pode ver que a China com 38,5% da capacidade global tem o maior

proporção, seguido pelos Estados Unidos com 16,1%. No geral, a capacidade instalada da WP aumentou substancialmente nos últimos 11 anos.

Atualmente, a geração de energia eólica tornou-se um componente importante da rede inteligente, microrredes inteligentes, edifícios inteligentes e casas inteligentes, e desempenha um papel importante no fornecimento de eletricidade. No entanto, devido à alta estocasticidade e baixa previsibilidade do vento, a energia eólica é uma fonte intermitente de energia, o que influencia muito a segurança e a estabilidade dos sistemas WP integrados à rede de larga escala [8]. Além disso, é afetado por perdas de transmissão e distribuição [9]. Portanto, programação, gerenciamento e otimização são os principais desafios para a alta penetração de fontes de energia renovável, incluindo energia eólica e solar em redes de diferentes escalas [10].

Uma das soluções promissoras para os desafios acima é prever com precisão o WP futuro. Tal previsão também desempenhará um papel ativo na redução de custos operacionais e aumentará a competitividade do WP [11, 12]. Portanto, a previsão de WP é uma tecnologia chave que deve ser estudada para integrar WP com redes elétricas existentes de diferentes escalas.

Na literatura, existem duas maneiras de realizar a previsão de WP. O primeiro é prever o WP diretamente com os dados históricos do WP, enquanto o segundo é prever primeiro a velocidade do vento (WS) e depois produzir

^{*} Autor correspondente.

Endereços de e-mail: wangyun19@csu.edu.cn (Y. Wang), rmzou@csu.edu.cn (R. Zou), liufang@csu.edu.cn (F. Liu), zhanglingjun@hdu.edu.cn (L. Zhang), liu7y@foxmail.com (Q. Liu).

<https://doi.org/10.1016/j.apenergy.2021.117766> Recebido

em 21 de abril de 2021; Recebido em formulário revisado em 24 de agosto de 2021; Aceito em 27 de agosto de 2021 Disponível online em 10 de setembro de 2021 0306-2619/© 2021 Elsevier Ltd. Todos os direitos reservados.

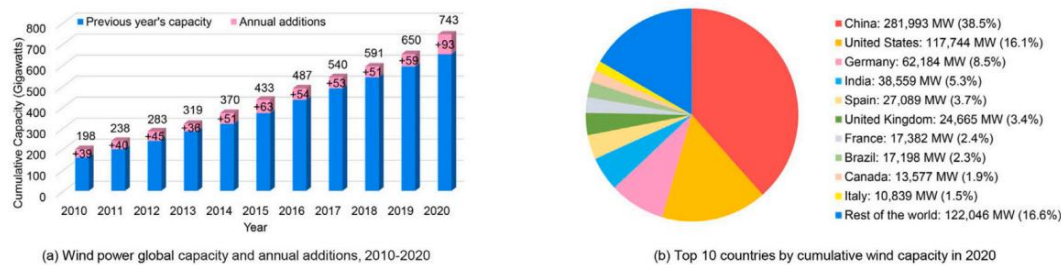


Fig. 1. Instalações globais de energia eólica.

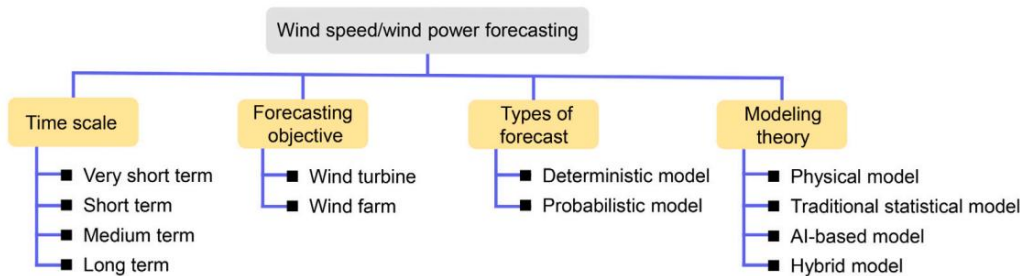


Fig. 2. Classificação da previsão de WS/WP.

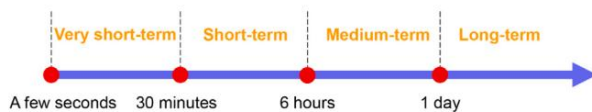


Fig. 3. Classificação dos métodos de previsão WS/WP com base em escalas de tempo.

WP prevê de acordo com uma curva de energia eólica [1,13], que reflete a energia que pode ser gerada em diferentes valores de WS. Portanto, os estudiosos prestaram atenção tanto à previsão de WS quanto à previsão de WP.

1.1. Técnicas de previsão de WS e WP

Muitas abordagens para previsão de WS e WP (WS/WP) foram propostas e implementadas. De acordo com diferentes padrões de classificação, eles podem ser divididos em vários grupos, conforme mostrado na Fig. 2.

Conforme mostrado na Fig. 3, os modelos de previsão WS/WP podem ser agrupados de acordo com suas escalas de tempo, incluindo *muito curto prazo* (alguns segundos a 30 minutos à frente), *curto prazo* (30 minutos a 6 horas à frente), *médio prazo* (6 h a 1 dia à frente) e modelos *de longo prazo* (1 dia a 1 semana ou mais à frente) [4,11,14]. Previsões de ultracurto prazo são úteis para controle de turbinas e rastreamento de carga em tempo real; previsões de curto prazo são usadas para planejamento de despacho de carga; as previsões de médio prazo são utilizadas para comercialização de energia e gestão do sistema de energia; previsões de longo prazo orientam a programação de manutenção (12,14).

De acordo com os diferentes objetivos de previsão, os modelos também podem ser classificados em dois tipos: *previsão WS/WP de turbina eólica e parque eólico* . O primeiro pode ajudar a caracterizar a potência de saída de uma única turbina eólica [8,15,16]. Este último integra informações de muitas turbinas eólicas para prever a potência de um parque eólico inteiro [17–20]. Das duas, a última abordagem é mais difícil.

Dependendo do tipo de previsão fornecida, os modelos WS/WP podem ser divididos em modelos *determinísticos* e *probabilísticos* [21]. Modelos determinísticos, que apenas fornecem previsões pontuais WS/WP [22–25], são limitados em seu desempenho de previsão devido à complexidade do ambiente e às vezes geram previsões insatisfatórias com muitos erros. Modelos probabilísticos, que muitas vezes fornecem mais informações para os tomadores de decisão do que previsões pontuais [22], podem representar incerteza em termos de intervalos de previsão (PIs) [19,22,24–26]. Na literatura, a previsão de incerteza pode ser dividida em

métodos paramétricos e não paramétricos [27]. Abordagens paramétricas geralmente assumem que a incerteza de previsão pode ser caracterizada por uma distribuição de probabilidade [28]. Modelos paramétricos representativos são o método delta, estimativa de variância média (MVE) e processo gaussiano (GP) [28]. O método delta assume que o termo do erro de previsão é distribuído gaussiano [29]. Os erros de previsão em modelos baseados em MVE obedecem à distribuição Gaussiana aditiva com variância não constante, enquanto toda combinação linear finita de variáveis aleatórias em GP é normalmente distribuída [30]. Em contraste, os métodos não paramétricos são livres de distribuição e visam gerar PIs para caracterizar a incerteza de previsão [28]. Dois populares são a estimativa do limite superior inferior (LUBE) [31] e a regressão quantílica (QR) [30]. Uma função de perda de Laplace assimétrica é empregada em modelos baseados em QR para ajudar a gerar PIs, enquanto uma medida de qualidade de PIs é geralmente adotada nos modelos baseados em LUBE.

Os modelos de previsão também podem ser divididos em quatro tipos com base nas diferenças na teoria da modelagem [25,32]: modelos *físicos* , modelos *estatísticos tradicionais* , modelos baseados em *inteligência artificial (IA)* e modelos *híbridos* .

Modelos físicos, como previsão numérica do tempo (NWP) e previsão de pesquisadores do tempo (WRF), geralmente consideram vários fatores meteorológicos (por exemplo, pressão do ar, umidade e temperatura) para prever WS futuro [12,25]. Estudos anteriores mostraram que os modelos físicos tiveram melhor desempenho na previsão de WS de médio e longo prazo [25], o que possibilitou obter previsões de WP usando a curva de potência esperada (12,14).

Modelos estatísticos tradicionais incluem média móvel autorregressiva (ARMA) [33], média móvel integrada autorregressiva (ARIMA) [34], ARIMA fracional (f-ARIMA) [35] e o modelo autorregressivo de Hammerstein [36]. Ao contrário dos modelos físicos, estes são modelos de séries temporais que normalmente são empregados para caracterizar a flutuação linear de WS/WP em diferentes locais e, em geral, apresentam bom desempenho na previsão de WS/WP de curtíssimo e curto prazo [12].

Com os avanços na ciência da computação, muitos modelos baseados em IA também têm sido amplamente utilizados na previsão de WS/WP, por exemplo, máquina de vetor de suporte (SVM) [37], máquina de vetor de suporte de mínimos quadrados (LSSVM) [38], máquina de aprendizado extremo (ELM) [39], kernel ELM (KELM) [40], método de lógica fuzzy [41], filtro de Kalman [42] e diferentes tipos de redes neurais artificiais (RNAs), incluindo rede neural de retropropagação (BPNN) [43], perceptron multicamada (MLP) [44], rede neural wavelet (WNN) [43], rede Elman [45] e regressão generalizada

rede neural (GRNN) [43]. Atualmente, com o desenvolvimento da tecnologia de aprendizado profundo, redes neurais profundas (DNNs), incluindo memória de longo prazo (LSTM) [39], rede neural convolucional (CNN) [19], LSTM bidirecional (BiLSTM) [46], unidade recorrente fechada (GRU) [47], GRU bidirecional (BiGRU) [48], rede de crença profunda (DBN) [49] e auto-encoder (AE) [50], atraíram maior atenção na previsão WS/WP devido à sua capacidade superior em lidar com problemas não lineares complexos. Em muitos estudos anteriores, os modelos baseados em IA mostraram melhor desempenho de previsão de WS/WP do que os modelos estatísticos tradicionais.

Modelos híbridos podem ser desenvolvidos combinando diferentes tipos de modelos para caracterizar diferentes aspectos da flutuação WS/WP [51]. Na literatura, os modelos híbridos de previsão são divididos em duas categorias com base em suas estratégias híbridas: modelos *baseados em empilhamento* e modelos *baseados em peso* [52]. Os modelos de previsão baseados em empilhamento geralmente consideram as previsões de um ou mais modelos de base como características, que são então combinadas com outro modelo de alto nível. Por exemplo, os recursos de saída da CNN foram alimentados em um LSTM para extrair recursos espaço-temporais dos dados de vento originais [19,53,54]. Em [55] e [56], vários LSTMs foram empilhados para formar um modelo híbrido para fazer pleno uso de informações espaço-temporais e múltiplos dados meteorológicos, respectivamente. Em [57], o SVM combinou as previsões de diferentes LSTMs para fazer previsões de WS.

No que diz respeito aos modelos de previsão baseados em peso, eles podem ser construídos com diversidade de previsão [58]. Por exemplo, em [44], três previsores de base diferentes, MLP, LSTM e ARIMA, foram combinados para prever WS, e os pesos de combinação correspondentes foram otimizados por um algoritmo de otimização inteligente. Em [43], quatro tipos diferentes de RNAs, BPNN, Elman, WNN e GRNN, foram agrupados usando certos pesos de combinação para fazer previsões de WS. Dois modelos híbridos (ARIMA-ANN e ARIMA-SVM) foram propostos para extrair simultaneamente flutuações lineares e não lineares em WS, e seu desempenho mostrou ser melhor do que os modelos de previsão simples [51].

Usar parâmetros de diversos previsores homogêneos é outra maneira de construir modelos baseados em peso. Em [57], seis LSTMs diversos foram selecionados como regressores de base para prever WS, enquanto em [59], nove AEs diferentes foram empregados como previsores de base.

O desempenho de diferentes abordagens de previsão é muito afetado pela definição da configuração do modelo, de modo que muitos pesquisadores adotam diferentes tipos de algoritmos de otimização inteligente para encontrar as configurações ideais de vários modelos de previsão WS/WP. Por exemplo, em [60], algoritmo genético (GA), busca de cuco (CS) e algoritmo de morcego gradiente conjugado (CG-BA) foram selecionados para ajustar três abordagens de previsão de WS, BPNN, WNN e GRNN, respectivamente. Em [61], os parâmetros ótimos do SVM foram ajustados por um algoritmo de busca atômica aprimorado (IASA), que teve um desempenho melhor do que o SVM otimizado para otimização por enxame de partículas (PSO), GA e algoritmo de busca atômica (ASA). Em [38], um otimizador híbrido PSOGSA, que combinou PSO com algoritmo de busca gravitacional, mostrou uma velocidade de convergência mais rápida que PSO no ajuste do modelo de previsão WS LSSVM. Alguns algoritmos de otimização inteligente também foram empregados para ajustar DNNs. Em [62], os autores projetaram um sistema híbrido de previsão WS, no qual GA foi usado para otimizar LSTM, enquanto em [63], o algoritmo DE foi usado para LSTM para equilibrar sua complexidade. A configuração ótima do BiLSTM foi determinada por uma nova versão do algoritmo de distribuição normal generalizada (GNDG) para previsão de WS [46]. Em muitos estudos anteriores, métodos de otimização de objetivo único foram frequentemente usados para melhorar o desempenho de RNAs na previsão de WS/WP, mas eles não conseguiram alcançar alta precisão e forte estabilidade na previsão simultaneamente [64]. Portanto, muitos métodos de otimização inteligente multiobjetivo foram adotados para resolver este problema. Por exemplo, o algoritmo multiobjetivo seno cosseno (MOSCA) para otimizar WNN foi relatado para gerar previsões WS precisas e estáveis em [64]. Em [65], o algoritmo de busca de cuco multiobjetivo (MOCS) foi empregado para otimizar o filtro Elman e Volterra, enquanto a otimização multiobjetivo de lobo cinzento (MOGWO) foi selecionada para ELM em [66].

Como a maioria das técnicas de previsão são orientadas por dados e podem, portanto, aprender padrões de diferentes dados de vento automaticamente, elas têm uma ampla gama de aplicabilidade em diferentes escalas de grade, como a grade em massa, microrredes e grades de construção e domésticas. Em outras palavras, diferentes tipos de modelos podem ser usados para fazer previsões de WS/WP com os dados independentemente de sua fonte. No entanto, há uma pequena diferença para grades multiescala quando as previsões de WS são obtidas por meio de várias técnicas de previsão. Para redes em escala predial ou doméstica, as turbinas eólicas com potências baixas são geralmente instaladas devido a limitações de espaço, enquanto a rede em massa e as microrredes utilizam turbinas eólicas com altas potências. Portanto, diferentes curvas de energia eólica são necessárias para produzir previsões de WP para redes de diferentes escalas.

Muitos programas de software de previsão WP foram desenvolvidos na indústria eólica integrando várias técnicas de previsão. Por exemplo, Prediktor é um sistema de previsão WP físico desenvolvido pelo Laboratório Nacional Risø na Dinamarca [67]. WindForTM, anteriormente conhecido como o software Wind Power Prediction Tool, que foi desenvolvido pela Technical University of Denmark [68], é um sistema de previsão WP de autoaprendizagem e autocalibração baseado em uma combinação de modelos físicos e técnicas avançadas de IA [69]. 1. O sistema de previsão eWind foi desenvolvido por uma empresa americana, Truewind, combinando modelos físicos e modelos estatísticos [70]. A Ferramenta de Previsão de Energia Eólica Avançada foi desenvolvida pelo instituto de "Solare Energieversorgungstechnik" na Alemanha e usa ANNs e previsões meteorológicas para fazer previsões WP [71]. Aleasoft é um produto de projetos de pesquisa liderados pela Universidade Politécnica da Catalunha no campo da previsão de energia que combina técnicas estatísticas e baseadas em IA para prever WP [72]. DeepMind e Google aplicaram algoritmos baseados em IA a fazendas com 700 megawatts de capacidade de energia eólica no centro dos Estados Unidos, e os resultados mostraram que esses algoritmos aumentaram o valor econômico da energia eólica em cerca de 20% [73]. Conforme descrito acima, abordagens baseadas em IA e modelos híbridos têm sido amplamente utilizados para prever WP na indústria eólica. As abordagens baseadas em IA podem caracterizar as complexas flutuações não lineares da energia eólica. Abordagens físicas, que consideram as características climáticas e ambientais dos locais observados, podem simular mudanças climáticas que afetam a flutuação do vento. Para prever o WP em diferentes locais com precisão, os modelos híbridos que combinam modelos físicos e modelos baseados em IA são frequentemente usados em software de previsão. No entanto, técnicas baseadas em aprendizado profundo que têm desempenho superior em comparação com métodos convencionais baseados em IA não foram totalmente incorporadas na maioria dos softwares de previsão atuais. Em aplicações reais, o maior desafio é que os modelos convencionais baseados em IA têm capacidade limitada para descrever flutuações de WP complexas, resultando em precisão de previsão de WP insatisfatória.

Como resultado, muitos modelos avançados baseados em aprendizado profundo devem ser considerados para o desenvolvimento de novos softwares de previsão WP.

1.2. Motivação e contribuições da revisão

Nos últimos anos, várias revisões sobre a previsão de WS/WP de diferentes perspectivas foram publicadas. Jung e Broadwater [12] revisaram modelos de previsão físicos, estatísticos, de correlação espacial, regionais e híbridos WS/WP, bem como algumas técnicas de previsão probabilísticas. Marugán et al. [74] revisaram a utilização das RNAs tradicionais na previsão de WS/WP. Zendeheboudi et al. [75] fez uma revisão abrangente sobre a aplicação do SVM na previsão de energia solar e eólica. Oito estratégias para previsão de WS/WP multi-passo à frente foram revisadas por Wang et al. em [2]. Tascikaraoglu et al. [76], Xiao et al. [77], e Ren et al. [78] revisaram quatro aspectos principais das abordagens de previsão WS/WP híbridas, ou seja, modelos baseados em peso, pré-processamento, otimização e modelagem de erros de previsão. Liu et al. [79] revisaram as tecnologias de otimização multiobjetivo na previsão de WS/WP. Qian et al. [80] revisaram muitos modelos de previsão de energia eólica híbridos baseados em decomposição. Liu e Chen [4] revisaram sete aspectos dos métodos de processamento de dados na previsão de WS/WP, incluindo decomposição, seleção de recursos, extração de recursos, eliminação de ruídos, resíduos

tabela 1
Todas as abreviaturas usadas no texto.

Abreviação	Nome completo	Abreviação	Nome completo
IA	Inteligência artificial	LSSVM	Máquina de vetores de suporte aos mínimos quadrados
ARMA	Média móvel auto-regressiva	MUITO DE	Erro absoluto médio
ARIMA	Média móvel integrada auto-regressiva	MAPA	Erro percentual absoluto médio
RNAs	Redes neurais artificiais	LSTM	Memória de curto prazo longa
EA	Codificador automático	LGBT	Máquina de aumento de gradiente de luz
COMO UM	Algoritmo de busca atômica	MVE	Estimativa de variância média
AVMD	Decomposição de modo variacional adaptativo	MLP	Perceptron de várias camadas
ACF	Análise de autocorrelação	MLR	Regressão linear de várias respostas
ACO	Otimização de colônia de formigas	MOPSO	Otimização de enxame de partículas multiobjetivo
AWDO-MSA	Combinação de otimização adaptativa conduzida pelo vento e	MHHOGWO	Combinação de otimização Harris Hawks e otimizador de lobo cinzento
	recozimento simulado		
BPNN	Rede neural de retropropagação	MOSCA	Algoritmo multiobjetivo seno cosseno
BiLSTM	LSTM bidirecional	MOCS	Algoritmo de busca de cuco multiobjetivo
BiGRU	GRU bidirecional	MOGWO	Otimização de lobo cinzento multiobjetivo
BGA	Algoritmos genéticos com código binário	MMODA	Algoritmo de libélula multiobjetivo modificado
CS	Pesquisa de cuco	MOBA	Algoritmo de morcego multiobjetivo
CG-BA	Algoritmo de morcego gradiente conjugado	MOMVO	Otimização multi-objetivo multi-universo
CEEMD	Concluir EEMD	MIM	Informação mútua
CEEMDAN	CEEMD com ruído adaptativo	PREÇO	Rede de entradas exógenas
CRBM	RBM contínuo	NWP	Previsão numérica do tempo
CNN	Rede neural convolucional	NCULSTM	Nova atualização de célula LSTM
CCN	Rede convolucional causal	OVMD	Decomposição de modo variacional ideal
ConvGRU	Rede de unidade recorrente fechada convolucional	ORELM	ELM robusto atípico
ConvLSTM	Memória convolucional de longo prazo	Ruim	Intervalos de previsão
CMI	MI condicional	PSO	Otimização de enxame de partículas
OSC	Otimização de enxame de frango	PSOGSA	PSO com algoritmo de busca gravitacional
CSNN	Rede neural de pico convolucional	PSR	Reconstrução do espaço de fase
CBP	Cascata BPNN	PMI	MI parcial
CSA	Algoritmo cruzado	PSOSA	PSO baseado em recozimento simulado
CG-BA	Algoritmo de morcego gradiente conjugado	PACF	Função de autocorrelação parcial
DNNs	Redes neurais profundas	PCA	Análise do componente principal
DBN	Rede de crenças profundas	QR	Regressão quantílica
DIAS	Denoising AE	QPSO	Otimização de enxame de partículas com comportamento quântico
EVOLUÇÃO	Evolução diferencial	RBM	Máquina de Boltzmann restrita
DE-HHO	Otimização híbrida DE-harris Hawks	RBFNN	Rede neural de função de base radial
OLMO	Máquina de aprendizado extremo	RNNs	Redes neurais recorrentes
EWT	Transformada wavelet empírica	RF	Floresta aleatória
EMD	Decomposição em modo empírico	RKELM	ELM de kernel reduzido
EEMD	Conjunto EMD	RFE	Eliminação de recursos recursivos
ESN	Rede de estado de eco	R2	Coefficiente de determinação
OE	Algoritmo de otimização extrema	RMSE	Erro quadrático médio
FA	Algoritmo do vaga-lume	SSA	Análise de espectro singular
FEEMD	EEMD rápido	SD	Decomposição secundária
f-ARIMA	Fracionado-ARIMA	SAE	Codificador automático empilhado
GP	processo gaussiano	SSAE	AE esparsa empilhada
GRU	Unidade recorrente fechada	SDAE	Codificador automático de remoção de ruído empilhado
GRNN	Rede neural de regressão generalizada	Seq2Seq	Sequência a sequência
GA	Algoritmo genético	SADE	Evolução diferencial autoadaptável
GCT	Teste de causalidade de Granger	SVM	Máquina de vetores de suporte
GOA	Algoritmo de otimização do gafanhoto	SMA	Algoritmo de molde de limo
GND0	Algoritmo de distribuição normal generalizada	SMAP	Erro percentual absoluto médio simétrico
EXCELENTE	Otimizador de lobo cinzento	VMD	Decomposição em modo variacional
LEME	Máquina de aprendizado extremo histerético	WS	Velocidade do vento
Fora	Algoritmo de pesquisa atômica aprimorado	WP	Força do vento
ICEMDAN	CEEMDAN melhorado	WS/WP	Velocidade do vento e energia eólica
AGORA	Análise de espectro singular aprimorada	WD	Decomposição Wavelet
IPCA	Análise aprimorada de componentes principais	WPD	Decomposição de pacotes wavelet
IDMDN	Rede de densidade de mistura profunda aprimorada	WSTD	Denoising de limiar suave wavelet
IHWOSCA	GWO híbrido aprimorado e algoritmo de seno cosseno	WRF	Previsão do pesquisador do tempo
KELM	Kernel ELM	WNN	Rede neural wavelet
KPCA	Análise de componentes principais do kernel	1D-CNN	CNN unidimensional
LUBRIFICANTE	Estimativa do limite superior inferior	2D-CNN	CNN bidimensional
LR	Regressão linear	3D-CNN	CNN tridimensional

modelagem de erros, detecção de outliers e correção baseada em filtros. Ahmed e Khalid [81] apresentou uma visão geral dos modelos de previsão para geração de energia renovável a partir de energia eólica e solar. Okumus e Dinler [82] apresentou o status atual dos modelos híbridos em WS/WP previsão e propôs uma abordagem híbrida para fazer uma hora de antecedência Previsão WS. Liu et al. [21] pesquisaram quatro tipos de modelos tradicionais baseados em IA e quatro tipos de DNNs na previsão de WS/WP, bem como duas estratégias auxiliares comumente usadas para modelos inteligentes híbridos Ahmad et al. [83] revisaram três tipos de previsão de energia eólica

modelos para a rede inteligente e edifícios, incluindo algoritmos de aprendizado de máquina, abordagens baseadas em conjuntos e RNAs. Khalid e Javaid [84] pesquisou vários algoritmos de otimização para previsão modelos para a rede inteligente, como carga, preço e previsão de energia eólica. Ssekulima et al. [85] revisaram dados físicos, estatísticos e baseados em IA metodologias de previsão de velocidade do vento e irradiância solar. Atualmente, juntamente com o desenvolvimento da informática e tecnologia, muitos tipos de DNNs foram desenvolvidos para diferentes propósitos. Eles deram uma grande contribuição para o desenvolvimento e utilização de energia eólica, especialmente na previsão de WS/WP [16,45,

47,86-95]. Mais de 100 estudos baseados em DNN para previsão de WS/WP foram publicados. Pela descrição do parágrafo anterior, é claro que os artigos de revisão anteriores prestaram pouca atenção às DNNs em diferentes aspectos da previsão de WS/WP. Assim, a principal contribuição deste artigo é revisar de forma abrangente o WS/WP baseado em DNN abordagens de previsão. As contribuições são resumidas a seguir.

- Este artigo revisa muitos estudos anteriores de acordo com os papéis de diferentes DNNs no processo de previsão de WS/WP, nomeadamente pré-processamento de dados, extração de recursos e aprendizado de relacionamento.
- Vantagens e desvantagens de diferentes DNNs utilizadas no mesmo processo de previsão WS/WP são comparados e discutidos de vários aspectos, principalmente em seu custo computacional.
- Além de modelos profundos únicos, DNNs híbridos para aprendizado de relacionamento também são revistos com base em diferentes estratégias híbridas, nomeadamente estratégia híbrida baseada em empilhamento e estratégia híbrida baseada em peso.
- Este artigo examina diferentes tipos de otimizadores para determinar a configuração ideal de vários modelos de previsão WS/WP, incluindo modelos baseados em kernel, modelos baseados em ANN e modelos híbridos baseados em peso.
- Este artigo discute os desafios de diferentes tipos de DNNs na previsão de WS/WP, e resume algumas pesquisas futuras instruções para melhorar ainda mais a precisão da previsão com DNNs.

Para os pesquisadores, este artigo pode fornecer conhecimento da teoria, status atual, desafios e tendências futuras de DNNs na previsão de WS/WP, que podem orientá-los no desenvolvimento de modelos de previsão mais precisos e eficientes usando DNNs adequados. Além disso, esperamos promover DNNs subvalorizadas para que possam ser amplamente testadas para Previsão de WS/WP. Para desenvolvedores de software, este documento fornece uma visão geral de muitos modelos avançados de aprendizado profundo que podem ser em consideração para melhorar a precisão da previsão de WP.

1.3. Organização do papel

O resto do artigo está estruturado da seguinte forma. Primeiro, a Seção 2 se concentra em resumir o processo de previsão de WS/WP e suas inerentes problemas. Em seguida, as Seções 3, 4, 5 e 6 são dedicadas à revisão os métodos usados nas quatro etapas principais da previsão de WS/WP com DNNs: processamento de dados, extração de recursos, aprendizado de relacionamento e otimização do modelo, respectivamente. A Seção 7 apresenta os resultados de estudos de caso numéricos para diferentes métodos. A Seção 8 descreve o desafios e tendências futuras na previsão de WS/WP. Por fim, a Seção 9 oferece conclusões provisórias do artigo. A Tabela 1 lista todas as abreviações usado neste papel.

2. Descrição da tarefa de previsão de WS/WP

O vento é produzido pelo movimento do ar e é afetado por muitos fatores meteorológicos [78]. Assim, na previsão de WS, além do histórico dados WS, alguns fatores meteorológicos, incluindo temperatura, pressão, e umidade, também são consideradas variáveis explicativas para caracterizar a flutuação da WS [96]. WP é produzido pela turbina eólica, o que significa que o vento é desviado pelas pás de uma turbina eólica e, em seguida, convertida em energia mecânica através da rotação, que aciona um gerador para produzir eletricidade e converter energia [78,97]. O poder desempenho de geração de uma turbina eólica pode ser descrito por um curva de potência, que mostra a relação entre a saída da turbina power e WS com a seguinte função [97],

$$P = \begin{cases} 0 & v < v_{in} \\ \frac{P_r}{v_r} (v - v_{in})^3 & v_{in} \leq v \leq v_r \\ 0 & v > v_r \end{cases} \quad (1)$$

onde P é a potência de saída da turbina em WS, v é o WP avaliado, P_r é a densidade do ar, A é a área de varredura de um impulsor, C_p é o WP coeficiente e denotam o WS de entrada, nominal e de corte,

respectivamente. Pode ser visto a partir da Eq. (1) que o WP é muito afetado pelo WS e densidade do ar, que muda com o clima. Portanto, além de dados históricos de WP, WS e outros fatores meteorológicos podem também ajudam a prever o futuro WP [88].

Neste artigo, uma série temporal WS ou WP é expressa como $\{ (1), (2), \dots, (\tilde{y}(1)), (\tilde{y}(2)), \dots, (\tilde{y}(n)) \}$, onde $(\tilde{y}(t))$ é o valor de WS/WP no tempo t , e n é o comprimento dos dados. Outras variáveis relacionadas ao flutuação de WS/WP são denotados como $1, 2, \dots, e$, onde e é o número total de variáveis, $e = \{ (1), (2), \dots, (\tilde{y}(1)), (\tilde{y}(2)), \dots, (\tilde{y}(n)) \}$ é a série temporal da a variável. Se $e = 0$, significa que nenhuma variável relacionada seja coletada.

Para fazer uma previsão de WS/WP \tilde{y} -passo à frente no momento, uma previsão explícita ou função implícita deve ser usada. Sua expressão é dada por

$$(\tilde{y} + \tilde{y}) = [(\tilde{y}), 1(\tilde{y}), \dots, (\tilde{y})] + \quad (2)$$

onde (\tilde{y}) , (\tilde{y}) são vetores de características que contribuem para a flutuação em $(\tilde{y} + \tilde{y})$, função (\tilde{y}) descreve a relação entre os recursos e meta de previsão com o conjunto de parâmetros do modelo e, denota o erro de previsão.

Para estimar o erro entre os dados reais e os correspondentes previsão, uma função de perda é geralmente definida [98]. O mais popular é a função de perda de mínimos quadrados, que é expressa como

$$\min_{\tilde{y}} \sum_{i=1}^n [(\tilde{y} + \tilde{y}) - \tilde{y}(\tilde{y} + \tilde{y})]^2 \quad (3)$$

onde $\tilde{y}(\tilde{y} + \tilde{y})$ é a previsão. Dada a configuração do modelo, a melhor O conjunto de parâmetros do modelo pode ser obtido minimizando a função de perda.

A expressão do modelo de previsão WS/WP mostrado na Eq. (2) mostra que a precisão dos resultados da previsão pode ser afetada por os quatro fatores a seguir:

- Qualidade dos dados de vento coletados. Devido a fatores estocásticos como um ambiente complexo e mutável e falhas de equipamentos, os dados de vento coletados podem conter várias incertezas, como valores ausentes, ruído e valores atípicos. A qualidade de $(\tilde{y} + \tilde{y})$, $1(\tilde{y})$, \dots , (\tilde{y}) na Eq. (2) pode ser muito afetado, impedindo-nos obter recursos adequados e treinar previsões precisas modelos. Portanto, antes de realizar a previsão WS/WP, é essencial pré-processar os dados de vento coletados para reduzir incerteza.
- Racionalidade dos recursos de entrada. Para um modelo de previsão, usando diferentes entradas (\tilde{y}) , $1(\tilde{y})$, \dots , (\tilde{y}) resultarão em diferentes desempenhos de previsão. Embora a melhoria na coleta de dados técnicas permitem que um número cada vez maior de recursos seja usado para prever futuros WS/WP, certos recursos redundantes ainda podem diminuir a capacidade de previsão dos modelos. Portanto, usando o apropriado recursos de entrada não apenas contribuem para a previsão precisa de WS/WP mas também torna o treinamento de modelos mais eficiente. Extrair útil características torna-se cada vez mais importante na previsão de WS/WP.
- Adequação do modelo de previsão. Devido à variabilidade climática, as características das flutuações WS/WP são tão complexas que um modelo simples não consegue descrevê-los de forma abrangente. Então, o modelo de previsão selecionado (\tilde{y}) deve ser capaz de modelar o relação não linear complexa entre os recursos de entrada e a previsão WS/WP. Além disso, de acordo com as características o erro de previsão, uma função de perda ótima pode ser projetada para fornecer os parâmetros ótimos do modelo. Previsão adequada modelos com parâmetros ideais contribuem para WS/WP precisos previsões.
- Configuração do modelo de previsão. Embora algumas previsões modelos podem abordar funções arbitrariamente complexas em teoria, a configuração do modelo tem um impacto maior na previsão desempenho na prática, como o tipo de função do kernel e os parâmetros de kernel correspondentes na previsão baseada em kernel modelos (por exemplo, SVM e LSSVM), e o número de camadas e nós ocultos, taxa de aprendizado e escolha de otimizador no tradicional

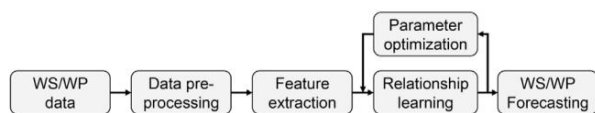


Fig. 4. Fluxograma de previsão de WS/WP.

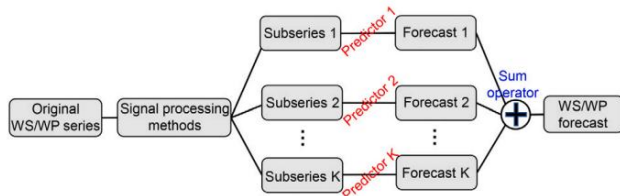


Fig. 5. Previsão de WS/WP com métodos de processamento de sinal.

redes neurais e DNNs. A configuração do modelo adequado pode guiar o modelo para aprender os parâmetros apropriados do modelo, de modo a obter um melhor desempenho de previsão.

Com base na análise acima, para obter previsões satisfatórias de WS/WP, os pesquisadores geralmente constroem modelos de previsão em quatro etapas principais: pré-processamento de dados, extração de recursos, aprendizado de relacionamento e otimização de parâmetros. Um fluxograma de toda a previsão WS/WP processo é ilustrado na Fig. 4.

Como os dados coletados podem conter ruídos desconhecidos ou valores discrepantes, os dados originais de WS/WP devem primeiro ser pré-processados por um método como um processamento de sinal ou detecção de outlier [4,99]. Então, características importantes podem ser extraídas para prever futuros WS/WP. Com os recursos selecionados, modelos de previsão WS/WP são implementados e treinados para aprender a relação entre os recursos de entrada e o alvo WS/WP. Posteriormente, os parâmetros ótimos dos modelos implementados devem ser determinados de acordo com sua previsão. desempenho nos dados fornecidos. Finalmente, com base na relação aprendida com a configuração ideal do modelo, as previsões WS/WP podem ser obtidos com os modelos treinados.

Com suas poderosas habilidades na extração e montagem automática de recursos de relacionamentos não lineares complexos, o aprendizado profundo tem sido amplamente aplicado no desenvolvimento e utilização de energia renovável, especialmente no processo de previsão de WS/WP. As quatro seções a seguir revisar as técnicas atuais usadas nas quatro principais etapas do WS/WP previsão com foco em aprendizado profundo.

3. Pré-processamento de dados de vento

Muitos métodos têm sido propostos e usados para lidar com incerteza nos dados de vento. Eles podem ser divididos basicamente em dois grupos: métodos de processamento de sinal e métodos de detecção de outliers [100,101].

Uma vez que uma série temporal WS/WP coletada pode ser considerada como um sinal, muitos pesquisadores usaram métodos de processamento de sinal para decompor dados de vento em diferentes subséries. Então, os resultados previstos de todos os subséries são agregadas para produzir previsões com base no original dados de vento [102]. O processo correspondente é mostrado na Fig. 5. Os métodos de processamento de sinal também podem ser usados como métodos de remoção de ruído para remover o ruído da série temporal WS/WP original, que produz os dados de vento sem ruído que são usados para treinar modelos de previsão [94, 103]. Atualmente, três tipos de métodos de processamento de sinal são usualmente usado para pré-processar séries temporais WS/WP brutas: abordagens baseadas em wavelets, métodos baseados em decomposição de modo e análise de espectro singular modelos baseados em (SSA).

As abordagens baseadas em wavelets incluem decomposição wavelet (WD) [104], decomposição de pacotes wavelet (WPD) [105-107], empírico

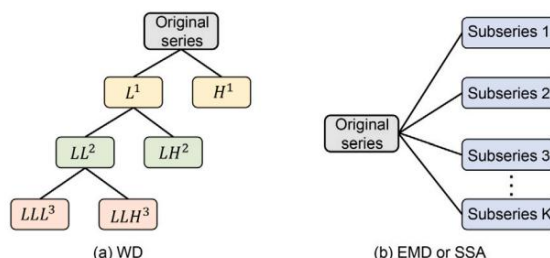


Fig. 6. Vários tipos de métodos de processamento de sinal.

mesa 2

Algoritmos SD existentes.

Artigo	método SD	Objetivo de processamento
Mi et al. [125]	WPD + EMD	WS
Xiang e outros [124]	SSA + VMD	WS
Wu et al. [126]	VMD + EMD, VMD + EWT WP	
	SSA + VMD, EMD + VMD	
	EWT + EMD, EWT + VMD	
Wu e Xiao [108]	VMD + SSA, EMD + SSA	WS
	EWT + SSA, SSA + EWT	
	EMD + EWT, VMD + EWT	
	SSA + EMD, VMD + EMD	
Moreno et al. [111,127]	VMD + SSA	WS
Zhang et al. [107]	VMD + WPD	WS
Sun et al. [112]	EEMD + AVMD	WS
Mi et al. [114]	SSA + EMD	WS
Qu et al. [116]	CEEMDAN + EWT	WS
Peng et al. [128]	CEEMDAN + VMD	WS
Liu et al. [129]	EEMD + WPD	WS

transformada wavelet (EWT) [108,109] e wavelet soft threshold de noising (WSTD) [94].

Os métodos baseados em decomposição de modo incluem decomposição de modo variacional (VMD) [107,110,111], VMD adaptativo (AVMD) [112], VMD ótimo (OVMD) [113], empírico

modo de decomposição (EMD) [102,114], conjunto EMD (EEMD) [115], EEMD rápido (FEEMD) [105], EEMD completo (CEEMD) [106], CEEMD com ruído adaptativo (CEEMDAN) [116,117], e CEEMDAN melhorado (ICEEMDAN) [43,118]. Além disso, SSA [111,114,119-121] e melhorias SSA (ISSA) [122] foram empregados para processar dados originais de WS/WP antes da previsão.

A maior diferença entre os três tipos de métodos de pré-processamento de dados acima é que as abordagens baseadas em wavelets podem decompor os dados originais do WS/WP em diferentes níveis, enquanto os outros podem só produz todas as subséries em um nível, como mostrado na Fig. 6.

Além dos métodos de processamento de sinal único, alguns métodos de processamento de sinal combinado, como decomposição secundária (SD), foram propostos e mostram maior eficiência para previsão de WS/WP [123,124]. Liu et al. [123] formulou um algoritmo SD com FEEMD e WPD para processar dados do WS. Em seu estudo, o WPD foi usado para decompor o dados em diferentes subséries com várias frequências, e as subséries de alta frequência foram decompostas pelo FEEMD. O final

Os resultados de previsão do WS mostraram que o modelo baseado em SD executou melhor do que as abordagens de processamento de sinal único. Desde então, muitos pesquisadores desenvolveram e estudaram vários algoritmos SD para Processamento de dados WS/WP. Os algoritmos SD existentes são resumidos na Tabela 2.

Modelos treinados com dados de vento contaminados por valores discrepantes podem incorporar informações confusas que resultam em mau desempenho de previsão [4]. Alguns pesquisadores usaram métodos de detecção de outliers para reconhecer outliers óbvios nos dados de vento, que são então corrigidos ou removido do conjunto de dados [4]. Xiong e Wang [99] utilizaram uma ponderação algoritmo de mínimos quadrados para detectar outliers, que foram corrigidos por interpolação spline cúbica. Os dados de vento modificados foram usados para treinar diferentes modelos de previsão. Wang et al. [130] aplicaram SVM para reconhecer outliers e substituí-los pelos valores estimados gerados

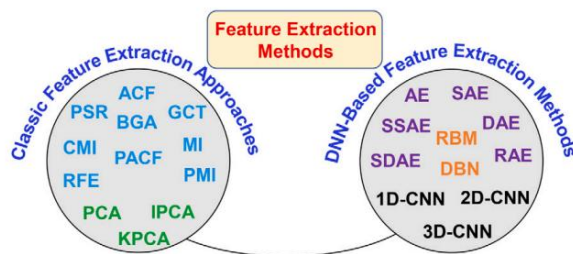


Fig. 7. Métodos de extração de características usados na previsão de WS/WP.

pela SVM. O critério de seis sigma foi utilizado por Ouyang et al. [131] para detectar outliers nos dados de energia eólica, enquanto o outlier baseado na densidade local método de detecção foi adotado em [132]. Ele et al. [133] empregado o método quartil para detectar outliers, que foram substituídos por valores estimado a partir da interpolação de spline cúbico antes de prever o WP.

Modelos baseados em aprendizado profundo, especialmente denoising AE (DAE), têm também foi utilizado para reduzir o ruído nos dados de vento originais. DAE, um variante de EA tradicional proposta por Vincent et al. [134], usa entrada reconstrução para gerar uma representação robusta de recursos na presença de ruído [98,134]. Embora nenhum estudo tenha afirmado que o objetivo da usando DAE na previsão WS/WP era obter dados de entrada sem ruído, foi empregado para extrair características robustas da entrada original dados, possibilitando reduzir ruídos e extrair recursos simultaneamente. Por exemplo, em [18,98,135–137], o DAE empilhado foi adotado para reduzir o ruído nos dados e extrair recursos robustos concomitantemente, seguido pelo uso de outros modelos para fazer previsões de WS/WP. Para evitar repetições desnecessárias, a descrição dos modelos baseados em DAE incluindo a teoria e aplicações na previsão de WS/WP foi apresentada na Seção 4.2.1, com subtítulo métodos de extração de características baseados em AE.

Infelizmente, ao revisar a literatura, poucos outros modelos baseados em deep learning foram capazes de pré-processar dados WS/WP e identificar outliers neles.

As diferenças entre os principais tipos de métodos de pré-processamento para os dados de vento podem ser resumidos da seguinte forma. Métodos de detecção de valores discrepantes para cus em lidar com amostras anormais nos dados de vento originais, enquanto métodos de processamento de sinal processam todos os dados indiscriminadamente [100]. Os métodos de processamento de sinal são geralmente usados para dois propósitos no pré-processamento de dados de vento: decomposição de dados e redução de ruído de dados, o que melhora a previsibilidade da subsérie decomposta por diferentes métodos de processamento de sinal é geralmente mais forte do que a dos dados originais, resultando em previsões precisas de cada parte dos dados originais. No entanto, remover dados ruidosos com métodos de processamento de sinal carrega o risco de que ainda existam valores discrepantes nos dados sem ruído [101]. Com outlier métodos de detecção, não sabemos se todos os outliers são detectados completamente [100]. Em geral, poucos estudos foram encontrados na literatura que empregam dois tipos de métodos de pré-processamento simultaneamente para reduzir a incerteza nos dados de vento.

4. Redes neurais profundas para extração de recursos

Um processo de seleção de recursos bem projetado desempenha um papel fundamental no WS/WP previsão [138]. Uma vez que as características apropriadas são favoráveis ao melhoria da precisão da previsão WS/WP, muitos pesquisadores têm tentou encontrar o maior número possível de recursos para explicar as flutuações em WS/WP. Ao mesmo tempo, eles tentaram encontrar expressões eficazes para os recursos disponíveis para reduzir os efeitos adversos da redundância informações sobre a fase de treinamento dos modelos de previsão. Na literatura, muitos métodos de extração de características têm sido usados para WS/WP previsão, que podem ser categorizadas em dois grupos: recurso clássico abordagens de extração e métodos de extração de recursos baseados em DNN, como mostrado na Fig. 7.

4.1. Abordagens clássicas de extração de recursos

As abordagens clássicas de extração de recursos podem ser agrupadas em duas categorias, a primeira das quais ajuda a selecionar recursos úteis de um conjunto de recursos candidatos, enquanto o segundo converte recursos candidatos em novas, alterando assim as formas das feições finais selecionadas.

Os métodos típicos da primeira categoria são a reconstrução do espaço de fase (PSR) [139,140], teste de causalidade de Granger (GCT) [138,141], função de autocorrelação (ACF) [138,141], função de autocorrelação parcial (PACF) [141,142], eliminação de recursos recursivos (RFE) [138], algoritmos genéticos codificados binários (BGA) [129], informação mútua (MI) [143, 144], MI condicional (CMI) [145,146], MI parcial (PMI) [147] e seleção de recursos relevantes de cima para baixo [148]. Na previsão de WS/WP, as abordagens representativas na segunda categoria incluem componentes principais análise (PCA) [138,149], PCA melhorado (IPCA) [150] e PCA do kernel (KPCA) [151,152].

Devido aos seus diferentes princípios, os métodos de extração de características podem extrair recursos que representam diferentes aspectos dos dados. Por exemplo, ACF e PACF ajudam a analisar as relações lineares entre diferentes séries temporais, enquanto KPCA pode produzir características não lineares. Portanto, alguns pesquisadores propuseram métodos híbridos para selecionar características ótimas para tirar vantagem de diferentes algoritmos de extração de características. Feng et al. [138] considerou PCA, GCT, ACF, PACF e RFE, e propôs um método de seleção de recursos híbridos para selecionar a entrada mais adequada variáveis para previsão de WS de curto prazo. Amjady et al. [144] proposto um método de seleção de recursos de dois estágios para selecionar recursos adequados para WP previsão. No primeiro estágio, o MI filtra características irrelevantes e, no segundo estágio, um critério de redundância detecta e remove características entre as entradas candidatas selecionadas [144]. Leng et al. [153] propôs um coeficiente de correlação linear baseado em MI para evitar e recursos redundantes, que apresentaram melhor desempenho do que o outros cinco modelos de seleção de recursos [153].

4.2. Métodos de extração de recursos baseados em DNN

Embora os métodos híbridos possam selecionar e extrair recursos melhores do que modelos únicos, nenhum deles pode lidar com as características profundas de modelos complexos. dados de vento com características altamente não lineares. DNNs como AE e A CNN, por outro lado, provou ser um método poderoso para extração de recursos não lineares em muitos domínios. Atualmente, baseado em EA métodos de extração de recursos, métodos de extração de recursos baseados em RBM e Métodos de extração de recursos baseados em CNN são amplamente usados em WS/WP previsão.

4.2.1. Métodos de extração de recursos baseados em AE

Os diferentes tipos de EAs que foram empregados para extração de recursos na previsão de WS/WP são elaborados nas seções a seguir com referência à literatura relevante.

(1) Codificador automático

A estrutura típica de um auto-encoder (AE), uma das mais DNNs populares usados em aprendizado não supervisionado, é ilustrado na Fig. 8, a partir do qual podemos ver que AE é uma rede neural totalmente conectada com três camadas: uma camada de entrada, uma camada oculta e uma camada de saída. A Fig. 8 mostra que ele compreende basicamente partes de codificador e decodificador. Dentro na parte do codificador, o vetor de entrada original \tilde{y} é mapeado para um representação \tilde{y} ($<$) com a função de ativação da camada oculta (\tilde{y}), que pode ser expresso como $\tilde{y} = \sum w_{ij} \tilde{y}_i$, onde w_{ij} é o peso, \tilde{y} é o viés, e \tilde{y} é frequentemente uma função sigmóide. Na parte do decodificador, o código é mapeado de volta para a função de recurso reconstruída (\hat{y}), com o $\hat{y} = \sum w_{ji} \tilde{y}_j + \tilde{y}_j$, onde w_{ji} são o peso e o viés, \tilde{y}_j é o viés, e \tilde{y}_j é o viés, respectivamente.

Um EA visa aprender uma nova representação dos dados originais por minimizando a diferença entre \tilde{y} e \hat{y} , que é medida pela erro quadrado [154], $\tilde{y}(\tilde{y}) = \tilde{y} - \hat{y}$. Além disso, um termo de regularização 2 é frequentemente adicionado para evitar overfitting.

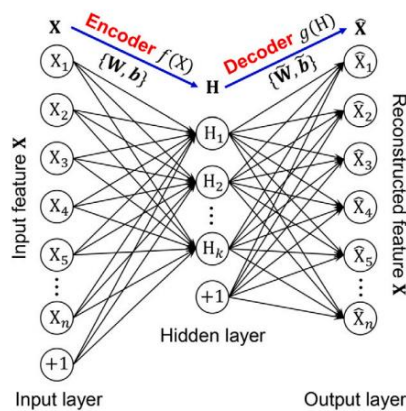


Fig. 8. Diagrama esquemático de um AE.

Atualmente, o uso da forma original de EA na previsão de WS/WP é raro. Mezaache e Bouzgu [50] usaram um EA para reduzir as dimensões de entrada para previsão de WS e uma rede neural Elman foi posteriormente usada para obter previsões. Os resultados mostraram que o modelo proposto teve um desempenho melhor do que um método baseado em ELM.

(2) Codificador automático empilhado

Um autoencoder empilhado (SAE) é uma rede neural multicamada profunda obtida pelo empilhamento de vários AEs [155–157]. A Fig. 9 mostra a estrutura de um SAE com três AEs. Para o primeiro AE, as entradas são os dados de vento originais. Para o i° AE subsequente, suas entradas são o código da camada oculta do $(i-1)^{\circ}$ AE.

Ao contrário do AE original, existem dois estágios principais para treinar um SAE: pré-treinamento não supervisionado em camadas e ajuste fino supervisionado [157]. Na primeira etapa, as características significativas são extraídas dos dados originais por vários EAs. Em seguida, as metas de previsão desejadas são empregadas como informações supervisionadas para ajustar os parâmetros da rede.

Usando um SAE, Jaseena e Kooor [156] propuseram um modelo híbrido baseado em DNN para prever WS. Em relação ao PCA, o SAE extrai recursos mais profundos e significativos do conjunto de dados original do WS. As características extraídas foram então usadas para treinar um modelo DNN para obter previsões. Jiao et al. [158] usaram um SAE com três camadas ocultas e uma função de perda subsequente para extrair as características da sequência de dados de referência. O número ótimo de neurônios na camada oculta e a taxa ótima de aprendizado de cada EA foram determinados pelo PSO. Liu et al. [159] empregaram um SAE com três camadas ocultas para aprender uma nova representação camada por camada no novo espaço de feição. Posteriormente, um SVM foi treinado com os recursos aprendidos. Os resultados de previsão do WS para Hong Kong demonstraram que o SAE pode fornecer um melhor espaço de recursos para dados de vento, e o modelo proposto mostrou melhor desempenho de previsão do que o SVM clássico.

(3) Codificador automático esparsamente empilhado

Um AE esparsamente empilhado, uma extensão de um AE, introduz um termo de penalidade esparsa na função objetivo original do AE [45,58]. Em um AE esparsamente empilhado, \tilde{y} é definido como o valor médio de ativação do i° nó da camada oculta e como o parâmetro esparsamente empilhado, que geralmente é definido como um valor pequeno próximo a 0. O termo de penalidade esparsamente empilhado pode ser expresso pelo método de Kullback-Leibler divergência de e :

$$(\tilde{y} | \tilde{y}) = \log \tilde{y} + (1 - \tilde{y}) \log (1 - \tilde{y}) \quad (4)$$

Com a restrição esparsa, um AE esparsamente empilhado pode produzir recursos esparsos e concisos, que podem expressar as características dos dados de entrada originais de forma mais adequada [160]. Um AE esparsamente empilhado (SSAE) tem uma estrutura semelhante à de um SAE, conforme mostrado na Fig. 9 [161]. A principal diferença entre eles é que diferentes AEs esparsamente empilhados no SSAE. Devido à sua capacidade de representar recursos esparsos, ele tem sido empregado para obter recursos adequados para previsão de WS/WP [45,58,59,162].

Qureshi et al. [59] usaram nove SSAEs como base-regressores, que foram treinados ajustando alguns parâmetros de aprendizado, incluindo o número de EAs empilhados, o número máximo de épocas e um termo de regularização de esparsidade. As previsões geradas pelos regressores de base treinados foram então alimentadas em um meta-regressor para obter as previsões finais do WP.

Em [58], Liu e Chen afirmaram que, além dos recursos gerados pelo método tradicional de transformação de resolução baseada em média, os recursos ocultos em dados WS de alta resolução também ajudaram a melhorar a precisão da previsão. Um SSAE de duas camadas foi aplicado para extrair a representação oculta dos dados originais do WS de alta resolução.

Em seguida, os recursos gerados a partir do método tradicional de transformação de resolução baseada em média e SSAE foram alimentados em DNNs para obter previsões de WS de 10 passos à frente.

Canção et al. [45] empregaram um MEEMD para decompor os dados WP em diferentes subséries. AEs esparsos foram usados para obter previsões para componentes de alta frequência e redes neurais de Elman para componentes de baixa frequência. Em seguida, as previsões finais do WP foram obtidas sobrepondo os dados anteriores e analisando o erro. Essa abordagem mostrou melhor desempenho na previsão de WP de ultracurto prazo do que o modelo de previsão sem usar AE esparsa.

(4) Codificador automático de denoising empilhado

Tem sido apontado que um AE comum, que apenas minimiza a função de perda, é incapaz de garantir a extração de características úteis, enquanto um DAE pode alterar os critérios de reconstrução via denoising para aprender uma boa representação [91,137]. Um DAE é semelhante a um AE convencional, exceto que adiciona ruído aos dados de entrada [91]. A ideia central do DAE é aprender uma representação robusta da entrada original de dados corrompidos. Para gerar o conjunto de dados corrompidos, um processo de corrupção estocástico é usado para selecionar aleatoriamente e mascarar os pontos de dados originais, definindo-os como zero. Depois disso, o conjunto de dados de entrada corrompido é alimentado no AE para gerar os dados de entrada reconstruídos \tilde{y} . O modelo é treinado minimizando o erro de reconstrução entre os dados de entrada não corrompidos e \tilde{y} . O processo de DAE é mostrado na Fig. 10. Ao empilhar vários DAEs, um DAE empilhado (SDAE) é construído com uma estrutura semelhante à de um SAE, conforme mostrado na Fig. 9.

Yan et al. [18] utilizaram uma rede multi-para-multi de um SDAE para corrigir dados NWP coletados de vários parques eólicos na mesma região. As características mais correlacionadas de muitas variáveis de entrada podem ser extraídas em diferentes locais dentro do parque eólico alvo. Então, para melhorar a robustez de um único SDAE, um modelo de previsão de conjunto SDAE WP foi desenvolvido com base em vários SDAEs com vários recursos e parâmetros de modelo importados. Os resultados mostram que o modelo proposto teve um desempenho melhor do que outros modelos de previsão individuais.

Peng et al. [135] usaram um SDAE com três camadas ocultas ajustadas por um algoritmo de retropropagação para capturar as características não lineares e complexas dos dados de vento. Os números ótimos de neurônios nas camadas ocultas de cada DAE foram determinados pelo algoritmo Bat. O SDAE demonstrou resultados de previsão eficazes.

A topografia e a meteorologia afetam as flutuações do WS e podem ser úteis para a previsão do WS, além dos dados históricos do WS [98]. No entanto, a existência de recursos não correlacionados e redundantes aumenta a complexidade da modelagem e reduz a capacidade de previsão. Liu et al. [98] propuseram um método de extração de características em dois estágios para selecionar características efetivas de um conjunto de fatores candidatos por um MI; em seguida, as características intrínsecas das características selecionadas foram capturadas pelo SDAE. Os resultados da previsão do WS mostraram que os recursos obtidos do SDAE podem melhorar a precisão da previsão do WS.

Jahangir et al. [136] projetaram um método multimodal baseado em módulos de redução de ruído e previsão para previsão de WS de curto prazo. Na fase de denoising, para diferentes tipos de entradas, um AG foi empregado para determinar a configuração ótima dos SDAEs, os quais foram aplicados para gerar recursos apropriados com menos ruído. Uma rede neural grosseira foi usada para ajustar o SDAE, cujo resultado foi usado para previsão de WS.

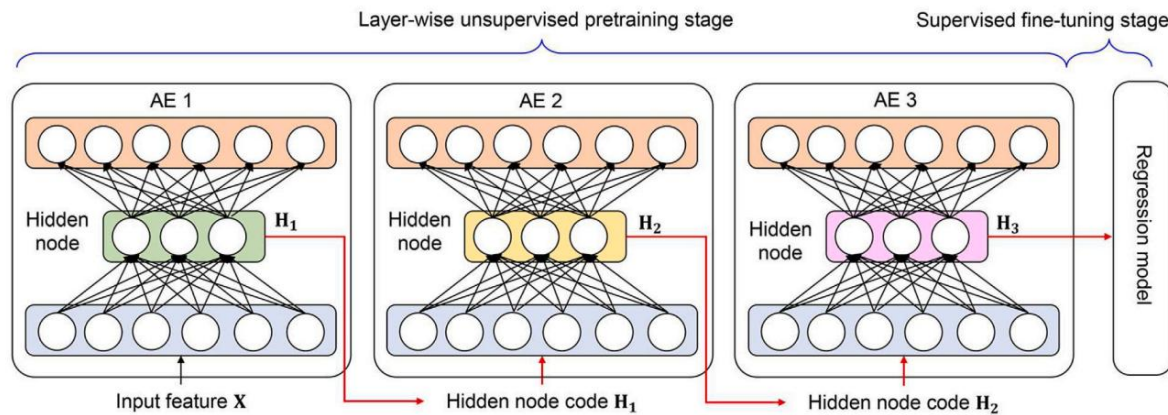


Fig. 9. Estrutura de um SAE com três EAs.

Em [137], um algoritmo de agrupamento de dados baseado em densidade foi usado para identificar outliers e selecionar amostras representativas. Em seguida, um SDAE com normalização em lote foi usado para extrair recursos profundos úteis. Os resultados mostraram que os recursos profundos extraídos melhoraram a precisão de previsão do WS do LSTM em 49% em comparação com a seleção de recursos sozinho. Isso sugere que a obtenção de recursos adequados é crucial para melhorar a precisão da previsão de WS.

(5) Codificador automático grosseiro

Em [155], Khodayar e Teshnehlab propuseram uma rede neural SAE com um neurônio bruto na camada de regressão para lidar com os fatores estocásticos em dados WS. Os resultados de previsão de WS de ultracurto e curto prazo no Colorado, EUA, mostraram melhorias promissoras em relação aos resultados de RNAs rasas e um SAE tradicional com uma camada de regressão convencional.

Para lidar com os fatores estocásticos em dados de vento, em [157], redes neurais brutas foram incorporadas ao AE e SDAE para desenvolver variantes brutas. Os resultados experimentais indicaram que o AE empilhado grosseiro proposto e o SDAE grosseiro poderiam gerar previsões WS de curto prazo com erro quadrático médio e erro médio absoluto mais baixos do que os métodos AE e SDAE convencionais e algumas redes neurais com arquiteturas rasas.

4.2.2. Modelos de extração de características

baseados em RBM Uma máquina de Boltzmann restrita (RBM), conforme mostrado na Fig. 11(a), contém duas camadas: camada *oculta* e camada *visível*. Um RBM permite conexões completas de unidades em diferentes camadas, mas não há conexões entre as unidades na mesma camada. Hinton apresentou um guia prático para treinar um RBM em [163]. DBNs são modelos de geração de probabilidade formados pelo empilhamento de vários RBMs. O processo de treinamento de um DBN pode ser dividido em duas fases semelhantes às dos modelos baseados em EA. A primeira fase é o pré-treinamento não supervisionado com um algoritmo de aprendizado guloso camada por camada. A principal tarefa da primeira fase é a otimização dos parâmetros iniciais do modelo, que são ajustados através do aprendizado supervisionado na segunda fase. Um diagrama esquemático de um DBN é mostrado na Fig. 11(b).

Em [164], uma DBN conhecida como *máquina de Boltzmann profunda* foi construída para fazer previsões de WS com antecedência de uma hora e de um dia, que mostrou uma melhoria de mais de 10% na precisão da previsão em comparação com um modelo autorregressivo linear, um neuro-fuzzy adaptativo sistema de inferência e um SVM. Em [165], além do WS histórico, vários outros atributos climáticos também foram levados em consideração para gerar previsões de WS com um DBN. Especificamente, um AG foi usado para otimizar os parâmetros no DBN e, em seguida, o DBN ajustado foi empregado para prever futuros WS em Taiwan. Um DBN também foi construído para previsão de WS em [49]. Em [166], ao contrário do DBN original, que foi formado pelo empilhamento de vários RBMs, Hu et al. empregou um RBM modificado chamado RBM contínuo (CRBM) para construir um novo DNN para lidar com

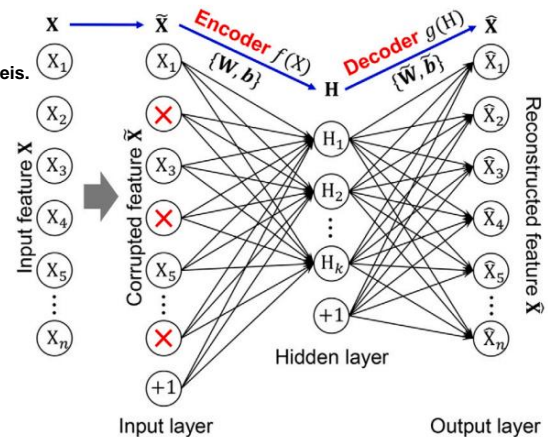


Fig. 10. Diagrama esquemático de um DAE.

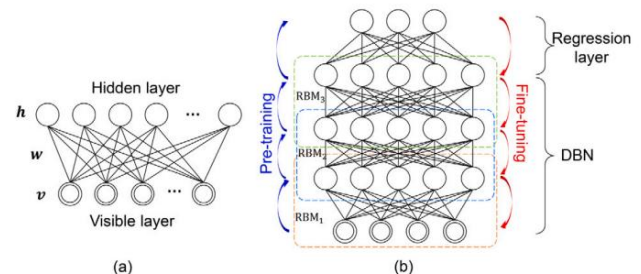


Fig. 11. Diagramas esquemáticos de (a) um RBM e (b) um DBN.

Problema de previsão do WS em Hong Kong. Comparado com o modelo baseado em EA empilhado, o modelo baseado em CRBM mais generativo produziu melhores previsões de WS [166].

Em [167], após a correção dos dados WP de acordo com os padrões relevantes, dados de três meses foram empregados para treinar um DBN com 3 camadas ocultas e 100 nós ocultos em cada camada para fazer previsões WP com 48 horas de antecedência. O DBN mostrou melhores resultados de previsão para a Mongólia Interior, China, do que as redes neurais existentes e um SVM. Em [168], dados de NWP (por exemplo, NWP WS, direção do vento, temperatura, umidade e pressão do ar), WS históricos e dados históricos de WP foram divididos em vários grupos por meio de agrupamento. Em seguida, amostras do cluster do dia de previsão foram selecionadas para treinar um DBN com uma camada de entrada, três camadas ocultas e uma camada de saída. Suas previsões para o parque eólico de Sotavento na Espanha superaram as de um BPNN e um Morlet WNN em > 44%.

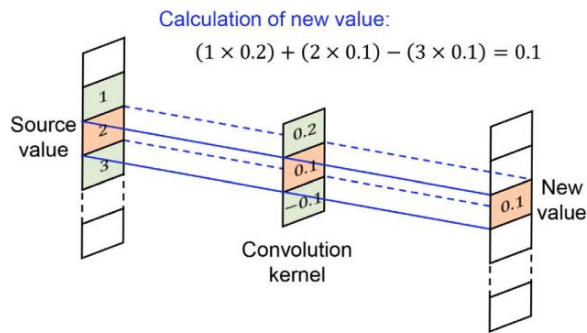


Fig. 12. Processo convolucional em 1D-CNN.

Wang et al. [22] usaram um DBN para desenvolver um modelo híbrido para previsão de WS determinística e probabilística. Inicialmente, a série temporal original do WS foi decomposta pelo WD em quatro sinais, incluindo um aproximado e três detalhados. Um DBN com duas camadas ocultas foi formado para prever o sinal aproximado, enquanto DBNs com oito camadas ocultas foram formulados para prever três sinais detalhados. As previsões determinísticas finais foram obtidas pela agregação das previsões das quatro sequências decompostas, e as previsões probabilísticas foram geradas por meio de regressão quantílica spline. A eficácia do modelo proposto foi validada para parques eólicos chineses e australianos.

Ele et al. [169] também desenvolveram um modelo híbrido para previsão de WS de ultracurto prazo. Um DBN foi empregado para extrair características de alta dimensão de quatro subséries geradas pelo WD. Ao contrário dos dois modelos híbridos acima, Zhang et al. [170] utilizaram SSA para decompor a série WP original em dois componentes: componentes de *tendência* e de *flutuação*. O primeiro foi previsto por um modelo LSSVM, enquanto um DBN foi empregado para prever a série flutuante, que apresentou maior aleatoriedade e incerteza. As previsões finais foram obtidas integrando as previsões dos componentes de tendência e flutuação. Da mesma forma, um DBN foi usado para prever uma subsérie WS em [122].

4.2.3. Métodos de extração de recursos baseados em CNN

A extração de recursos baseada em CNN é amplamente utilizada e eficaz em muitos campos, como processamento de imagens e reconhecimento de padrões [19, 171]. Uma CNN simples geralmente compreende quatro camadas: entrada, convolucional, pooling e camadas totalmente conectadas. Devido à sua poderosa capacidade de extrair recursos, diferentes modelos CNN têm sido usados para melhorar a precisão da previsão de WS/WP. Eles podem ser divididos em CNN unidimensional (1D-CNN), CNN bidimensional (2D-CNN) e CNN tridimensional (3D-CNN), cujos núcleos se movem em 1, 2 e 3 direções, respectivamente. Diferentes modelos CNN são usados para diferentes aplicações.

(1) CNN unidimensional Para

1D-CNN, tanto o kernel quanto o objetivo da pesquisa são geralmente vetores. No processo convolucional, como mostrado na Fig. 12, um elemento central do kernel é colocado sobre os dados de origem, que são então substituídos por uma soma ponderada de si mesmo e dos dados próximos. Através da operação de convolução, as feições locais dos dados originais podem ser extraídas e usadas para formar um conjunto de mapas de feições, cujos tamanhos são então reduzidos na camada de agrupamento por down-sampling, que agrega as feições em uma subárea selecionada de os mapas de características [19]. Geralmente, existem três métodos populares de downsampling: pooling máximo, pooling médio e pooling estocástico [172]. Finalmente, os mapas de recursos agrupados são transformados em um vetor 1D, que é então alimentado na camada totalmente conectada.

Hong et al. [173] utilizaram operações de convolução, kernel e pooling em um 1D-CNN para extrair características WP, que posteriormente foram alimentadas em uma rede neural de função de base radial supervisionada (RBFNN) com uma função de ativação gaussiana dupla para fazer 24 horas à frente Previsões de WP. Eles também estudaram os efeitos sobre o desempenho final da previsão

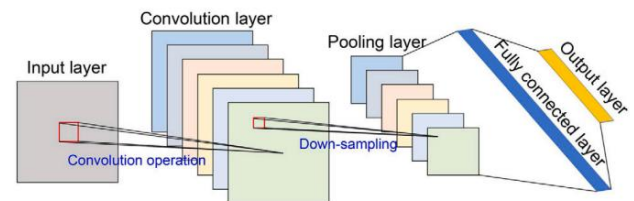


Fig. 13. Diagrama esquemático de um 2D-CNN.

das configurações do 1D-CNN, por exemplo, o número de camadas convolucionais/pooling, comprimento de kernels e taxa de dropout. Os resultados da previsão mostraram que em diferentes estações do ano, o modelo proposto superou todos os outros modelos que foram comparados.

Do ponto de vista físico, o WS compreende dois componentes: componentes médios e turbulentos [23]. Zhao et al. [23] desenvolveram um novo modelo baseado em dados para prever o WS médio e o desvio padrão da turbulência do vento. No modelo proposto, um 1D-CNN foi empregado para extrair informações acopladas ao tempo de dados históricos. Então, uma previsão determinística foi obtida através de uma camada totalmente conectada no 1D-CNN. A eficácia do modelo proposto foi validada com dados reais de um parque eólico chinês.

Em [174], Liu et al. propuseram um modelo híbrido de previsão de WS usando WPD para decompor os dados de WS em várias subséries com diferentes frequências. Posteriormente, um 1D-CNN foi empregado para prever as subséries com alta frequência. Os resultados mostraram que o modelo proposto, mesmo com mudanças bruscas de WS, teve um desempenho melhor que os demais modelos utilizados para comparação.

Harbola e Coors [90] projetaram duas novas redes para prever WS e direção do vento com um 1D-CNN baseado em um conjunto de dados de vento temporal. Uma única CNN 1D, a primeira rede das duas com três camadas convolucionais e duas camadas totalmente conectadas, usou os dados temporais de WS e direção do vento como entradas. O segundo foi um modelo 1D múltiplo de CNN combinando vários 1D-CNNs, mas com diferentes visualizações da mesma entrada. Assim, foi capaz de aprender mais recursos do que a única CNN 1D.

Ju et al. [89] investigaram as características das séries temporais WS de um local de estudo e seus adjacentes. Em seu estudo, as características originais de diferentes sites foram usadas em 1D-CNN para extrair informações úteis.

Em seguida, os dados da camada achatada do 1D-CNN foram alimentados em uma máquina de aumento de gradiente de luz (LGBM). Os resultados mostram que o modelo proposto foi superior ao SVM, LGBM e CNN em precisão e eficiência de previsão.

Rede convolucional causal (CCN) [175] é um tipo especial de CNN 1D. A saída convolucional de um CCN no tempo depende apenas das entradas antes do tempo e não tem nada a ver com entradas futuras [175].

Zhang e Liu [175] aplicaram um CCN para extrair características mais eficazes de diferentes subséries geradas por vários métodos de decomposição, o que melhorou a precisão da previsão.

(2) CNN bidimensional

Uma estrutura simples para 2D-CNN é mostrada na Fig. 13, e seu processo convolucional é mostrado na Fig. 14. Ao contrário de 1D-CNN, a entrada de 2D-CNN é uma matriz, e o kernel também é uma matriz. Como a entrada para um 2D-CNN deve ser um formato de matriz como uma imagem, muitos pesquisadores usaram métodos diferentes para formar entradas de matriz. Na literatura, esses métodos podem ser divididos em dois grupos de acordo com a diversidade dos dados de entrada.

No primeiro grupo, além dos dados originais de WS/WP, outras informações como fatores meteorológicos ou dados de WS/WP de locais adjacentes também podem ser coletadas, permitindo que a matriz de entrada seja construída a partir dos dados históricos para esses fatores, como mostrado na Fig. 15.

Em [19], informações espaciais e temporais coletadas de um parque eólico alvo e parques eólicos adjacentes foram usadas para construir gráficos de características sequenciais de tempo. Para extrair recursos úteis, os pesquisadores

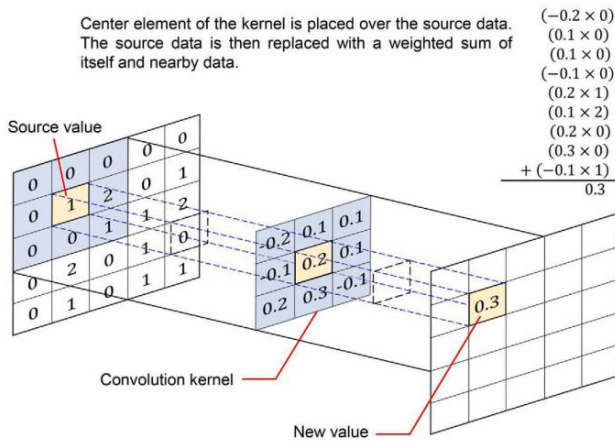


Fig. 14. Processo convolucional em 2D-CNN.

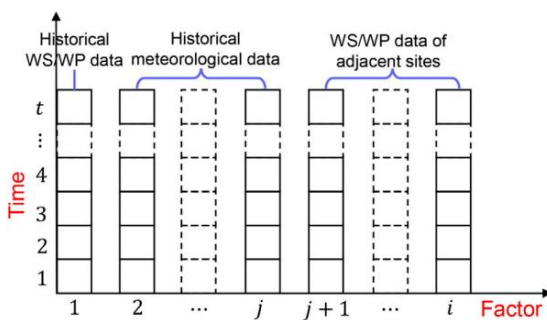


Fig. 15. Construção da matriz de entrada.

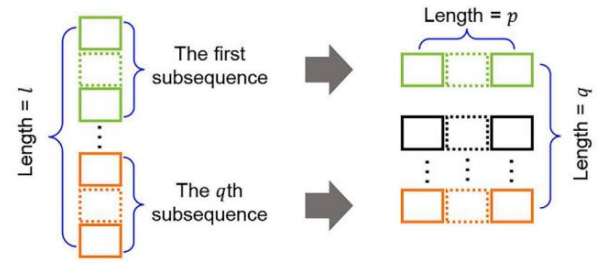


Fig. 16. Processo de transformação de dados 1D em dados 2D.

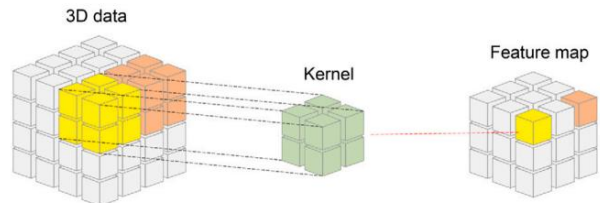


Fig. 17. Processo convolucional em um 3D-CNN.

utilizou uma rede neural híbrida, na qual as feições espaciais foram extraídas por um 2D-CNN e as feições temporais foram extraídas por um LSTM. Finalmente, QR foi empregado para fazer previsões probabilísticas de WP regional. Os resultados mostraram que o uso de informações espaço-temporais ajuda a melhorar a precisão da previsão.

Semelhante a [19], para fazer previsões de WS com 24 horas de antecedência em um parque eólico offshore, Hong et al. [176] usaram dados históricos do WS do local de estudo, além de informações auxiliares do WS de outros locais em Taiwan, China, Coréia e Filipinas. Esses dados históricos de WS espaço-temporais foram então alimentados em um 2D-CNN, cujos parâmetros estruturais e hiperparâmetros foram determinados pelo método de projeto robusto de Taguchi. Os resultados mostraram que os dados espaço-temporais do WS ajudaram a melhorar a precisão da previsão do WS, e usar o método Taguchi para determinar a estrutura e os hiperparâmetros do 2D-CNN foi melhor do que usar experimentos de tentativa e erro ou fatoriais completos.

Zhu et al. [53] estudaram as correlações espaciais e temporais de diferentes séries temporais de WS. Em seguida, um 2D-CNN foi empregado para extrair feições espaciais de matrizes WS espaciais, seguido por um LSTM para capturar as dependências temporais entre as feições espaciais. Dentro da estrutura proposta, o modelo aprendeu as correlações espaciais e temporais em conjunto e apresentou melhor desempenho do que outros seis modelos comparáveis.

Yu et al. [177] propuseram uma "série temporal de cenas" ao longo de um período de tempo para representar o processo de variação temporal e espacial. Essa série também pode ser considerada como uma imagem multicanal, e um 2D-CNN foi empregado para extrair as características espaciais e temporais e fazer previsões de WP. O modelo proposto alcançou uma precisão muito maior do que vários modelos de previsão WP existentes em diferentes horizontes de previsão.

No segundo grupo de métodos para criar a entrada da matriz, apenas os dados WS/WP originais são usados para construir as entradas da matriz, conforme mostrado na Fig. 16 [171,178]. O comprimento dos dados, denotado por

dividido em subsequências de comprimento . Cada subsequência é organizada sequencialmente como uma linha de uma matriz de entrada \times .

Wang et al. [178] transformaram dados 1D WP em uma matriz 2D usando reconstrução de dados. Em seguida, por meio de um 2D-CNN, a camada convolucional e a camada down-sampling foram usadas para extrair recursos ocultos dos dados do WP, enquanto uma camada de regressão logística foi utilizada para a previsão do WP. Com base no 2D-CNN, as previsões determinísticas e probabilísticas do WP foram obtidas por meio de uma abordagem de conjunto.

As condições climáticas têm sido relatadas como tendo um grande efeito nas mudanças no WS. Li et al. [96] utilizaram dois tipos de recursos para prever WS. Um vetor de recursos de entrada gerado por um 2D-CNN com dados históricos de WS foi selecionado como o primeiro tipo de entrada. Um segundo vetor de características compreendeu características meteorológicas ótimas, que foram obtidas por um MLP. Finalmente, os recursos extraídos foram alimentados em um modelo para prever WS. Dados do parque eólico na província de Gansu, na China, foram usados para testar a eficácia do modelo proposto.

Zhu et al. [171] investigou 400 entradas para fazer previsões de WP multi-passo à frente (Passo 1 ao Passo 16). Primeiro, as 400 entradas foram transformadas em uma matriz de 20×20 entradas. Em seguida, duas camadas convolucionais e duas camadas de agrupamento foram usadas para extrair recursos, que foram alimentados em um perceptron de camada única para prever o WP. A CNN 2D construída mostrou um bom desempenho de previsão WP para o parque eólico Elia na Bélgica.

(3) CNN tridimensional Ao

contrário das CNNs 1D e 2D, uma CNN 3D pode extrair características importantes dos dados de entrada 3D. Por exemplo, informações espaciais e temporais podem ser obtidas simultaneamente por meio de convolução 3D, cujo processo é representado na Fig. 17. Essa figura mostra que o kernel em uma CNN 3D também tem três dimensões. A movimentação do kernel através dos dados de entrada 3D permite obter um mapa de características 3D que contém informações que caracterizam as características de duas perspectivas diferentes, como as variações espaciais e temporais dos dados. Devido às suas vantagens, o 3D-CNN também tem sido usado para melhorar a precisão da previsão de WS/WP.

Higashiyama et al. [88] projetaram um 3D-CNN para extrair características espaciais temporais de dados NWP contendo várias variáveis meteorológicas previstas em uma grade horizontal de 5 km ao redor da área de um parque eólico alvo. Nesta fase, informações espaço-temporais úteis com menor dimensionalidade foram obtidas automaticamente, que juntamente com os dados históricos de WP foram usadas para treinar o preditor de WP. O modelo 3D baseado em CNN teve um desempenho muito melhor do que o método de previsão WP baseado em 2D-CNN [88]. Isso mostrou que um 3D-CNN ajuda a extrair

recursos mais úteis de dados NWP do que um 2D-CNN, resultando em uma previsão WP mais precisa.

Liu et al. [179] desenvolveram um modelo de previsão WS espaço-temporal baseado em uma rede de unidades recorrentes fechadas convolucionais (ConvGRU) e um 3D-CNN dentro de uma estrutura de previsão de codificação. Na fase de codificação, o 3D-CNN foi usado para extrair características de menor dimensão das saídas da camada ConvGRU, enquanto um 3D-CNN deconvolucional foi utilizado para expandir as saídas da camada deconvolucional ConvGRU na fase de previsão.

Mehrkanoon [180] usou 1D-, 2D- e 3D-CNNs para prever WS para várias estações meteorológicas. Diferentes características inerentes foram extraídas dos dados meteorológicos multivariados espaço-temporais. Os resultados mostraram que o modelo de previsão WS baseado em 3D-CNN teve um desempenho melhor do que o modelo baseado em 1D-CNN e o modelo baseado em 2D-CNN em 6 e 12 horas antecipação de WS [180]. Achados semelhantes com um 3D-CNN foram relatados em [181].

4.3. Comparação de diferentes modelos de extração de recursos

Conforme observado acima, os modelos clássicos de extração de recursos podem ser divididos em duas categorias. A primeira categoria apenas seleciona feições efetivas das feições originais (por exemplo, PACF e MI), enquanto a segunda categoria cria e condensa novas feições ou uma nova representação com base nas feições originais (por exemplo, PCA e KPCA). A maior diferença entre esses tipos de modelos é que o primeiro emprega a seleção supervisionada de recursos usando metas de previsão, enquanto o segundo extrai recursos de maneira não supervisionada. Quando há informações não confiáveis sobre as metas de previsão, os modelos não supervisionados são menos afetados; caso contrário, alvos de previsão adequados podem ser usados com modelos supervisionados para selecionar os recursos ideais.

Existem várias diferenças importantes entre os três tipos de abordagens de extração de características baseadas em DNN. No que diz respeito ao formato de entrada, modelos baseados em CNN podem extrair características de vetores, matrizes ou tensores, enquanto as entradas para modelos baseados em AE e modelos baseados em RBM devem ser apenas vetores, indicando que modelos baseados em CNN podem extrair mais tipos de recursos do que os outros dois tipos de modelos de extração de recursos. Em termos de processo de treinamento, modelos baseados em AE e modelos baseados em RBM geralmente têm duas fases: pré-treinamento *não supervisionado* em camadas e ajuste fino *supervisionado*, enquanto apenas um processo de treinamento supervisionado é empregado em modelos baseados em CNN. No caso de muitas camadas ocultas, é mais demorado extrair melhores características dos modelos baseados em AE e modelos baseados em RBM. Há também diferenças entre as funções objetivo de um único AE e RBM; o treinamento de EA minimiza o erro quadrático, enquanto uma função de energia é projetada e utilizada como objetivo de treinamento para RBM.

Uma comparação de diferentes tipos de CNNs é apresentada na Tabela 3. A principal diferença é a dimensionalidade dos dados de entrada e a maneira como o filtro desliza pelos dados. Para 1D-CNN, o kernel desliza apenas em uma direção, enquanto para 2D-CNN e 3D-CNN ele se move em duas e três direções, respectivamente. Portanto, 3D-CNN geralmente tem a vantagem de extrair recursos espaciais e temporais simultaneamente, enquanto 2D-CNN e 1D-CNN focam na captura de recursos espaciais e de padrões simples, respectivamente. Embora as informações extraídas por 3D-CNN sejam mais ricas do que as informações obtidas por 1D-CNN e 2D-CNN, a convolução 3D em 3D-CNN sofre da maior complexidade computacional, que é

() para entrada com dimensões de $x \times x$ com um $x \times x$ núcleo [182,183]. Assim, CNN 3D precisa de mais tempo computacional para treinamento do modelo do que 1D-CNN e 2D-CNN [176].

Embora os métodos de extração de recursos tradicionais e baseados em DNN possam obter novas expressões de recursos originais, há uma grande diferença entre eles. As abordagens tradicionais geralmente geram as novas expressões por algoritmos não supervisionados, enquanto todos os modelos de extração de recursos baseados em DNN têm fases de aprendizado supervisionadas. Como resultado, para diferentes tarefas de previsão com as mesmas entradas, as novas expressões de características obtidas pelos métodos tradicionais são as mesmas, mas para modelos baseados em DNN, elas variam com as tarefas de previsão

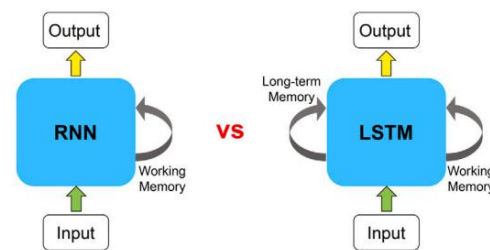


Fig. 18. Comparação da estrutura de RNN e LSTM.

metas de previsão. Além disso, os métodos tradicionais fornecem apenas uma representação de recursos rasos, enquanto recursos profundos podem ser representados por modelos baseados em DNN.

Diferentes tipos de modelos de extração de recursos são adequados para extrair informações apropriadas que refletem a flutuação complexa nos dados originais de WS ou WP. Em aplicações reais, as características da tarefa e os dados disponíveis determinam o tipo de métodos de extração de recursos usados.

5. Redes neurais profundas no aprendizado de relacionamento

Na prática, vários fatores afetam as variações de WS/WP. Dadas as relações complexas nos dados, os modelos tradicionais de regressão linear (por exemplo, ARMA, ARIMA ou um modelo persistente) e modelos de regressão não linear simples (por exemplo, SVM e BPNN) não caracterizam bem a complexidade. Portanto, muitos pesquisadores recorrem a DNNs com sua capacidade superior de ajuste não linear para tarefas de previsão de WS/WP. Nas subseções a seguir, a pesquisa sobre previsão de WS/WP é revisada com referência aos diferentes tipos de DNNs.

5.1. Memória de curto prazo longa e suas variantes

Aqui, apresentamos a técnica LSTM e suas variantes, bem como sua aplicação na previsão de WS/WP.

5.1.1. Memória de curto prazo longa

Como um problema de desaparecimento de gradiente pode ocorrer durante o processo de aprendizado de redes neurais recorrentes (RNNs), elas são incapazes de trabalhar com sequências mais longas e manter dependências de longo prazo, fazendo com que sofram de memória de curto prazo. LSTM, uma espécie de RNN, foi proposto por Schmidhuber e Hochreiter em 1997 para resolver este problema através de um mecanismo de gating [184]. Uma unidade LSTM comum é composta por uma célula, uma porta de entrada, uma porta de saída e uma porta de esquecimento [184]. A célula lembra as informações em intervalos de tempo arbitrários, e as três portas regulam o fluxo de informações para dentro e para fora da célula. As estruturas das redes RNN e LSTM são comparadas na Fig. 18, que mostra que LSTM usa memória de longo prazo adicional para lembrar informações passadas. As vantagens do LSTM o tornam amplamente utilizado para melhorar a precisão da previsão de WS/WP.

Devido à sua capacidade de previsão superior com séries temporais, o LSTM tem sido usado por muitos pesquisadores para obter diretamente previsões determinísticas ou probabilísticas de WS/WP com base em dados brutos de WS/WP. Por exemplo, em [24], o LSTM foi usado para gerar previsões WP determinísticas diretas. Os intervalos de previsão foram produzidos por uma distribuição beta ajustada para o erro de previsão correspondente. Banik et al. [26] desenvolveram um modelo baseado em LSTM para produzir intervalos de predição WP, nos quais os parâmetros do modelo foram ajustados por técnicas de otimização inteligente. Araya et al. [185] projetaram um modelo LSTM multi-escala para cada escala temporal e integrou suas informações de acordo para fazer previsões de WS de hora em hora para o dia seguinte. Yu et al. [186] desenvolveram uma rede de portão de esquecimento aprimorada por LSTM para previsão de WP.

Para melhorar a previsibilidade dos dados brutos de WS/WP, métodos de decomposição são frequentemente usados e o LSTM é geralmente selecionado para previsões. Em [143], diferentes modelos LSTM com estrutura otimizada via

Tabela 3

Uma comparação de diferentes tipos de CNNs.

Modelo	Direção	Características do recurso extraídas	Tamanho de entrada	Tamanho do kernel	Complexidade computacional
1D-CNN	Uma direção \hat{y}	Padrão simples	$\times 1$	$\times 1$)
2D-CNN	Duas direções $\hat{y} \hat{y}$	Temporal	\times	\times)
3D-CNN	Três direções $\hat{y} \hat{y} \hat{y}$	Espaço-temporal	$\times \times$	$\times \times$	(((((

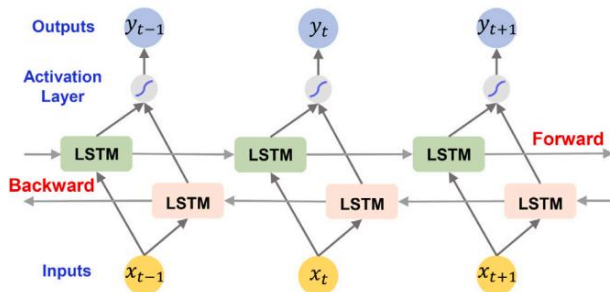


Fig. 19. Estrutura do BiLSTM com três etapas consecutivas.

o algoritmo de busca de corvo foi utilizado para prever subséries geradas pela WD, que mostrou que o LSTM ajustado teve melhor desempenho do que o LSTM original. Muitos estudos também usaram o LSTM para modelar Subsérie WS/WP [62,187–189]. No entanto, em [39], diferentes modelos foram empregados para prever subséries WS com diferentes frequências. Especificamente, o LSTM foi usado para prever uma subsérie de baixa frequência, enquanto o ELM foi usado para uma subsérie de alta frequência.

Depois que os recursos são extraídos usando um método como um método baseado em CNN, modelo ou modelo baseado em EA, os recursos são alimentados em um LSTM para produzir previsões. Em [98], dentro de um processo de aprendizado supervisionado SDAE, O LSTM foi empregado para gerar previsões de WS. Os resultados experimentais demonstraram que o modelo proposto teve um desempenho melhor do que o SAE e LSTM. Uma utilização semelhante do LSTM pode ser encontrada em [137]. Em [96] e [54], foram desenvolvidos modelos híbridos para fazer previsões de WS, em qual LSTM foi usado para capturar as dependências temporais entre os recursos extraídos usando CNN.

Tem sido relatado que os modelos de previsão combinados podem levar vantagem de modelos individuais e, portanto, geralmente apresentam melhor desempenho. Em [190], um LSTM, um ELM e uma rede neural Elman foram selecionados como preditores individuais para prever WS e, em seguida, mecanismos de combinação baseados em ELM foram usados para obter as previsões finais. Em [63], uma combinação não linear foi alcançada com um LSTM otimizado por algoritmo de evolução diferencial. A estratégia proposta mostrou melhor Resultados de previsão de WS do que linear, SVM e não linear baseado em ANN métodos de combinação.

5.1.2. Memória bidirecional longa e curta

Ao contrário do LSTM, o BiLSTM processa dados sequenciais em direções para trás com duas camadas ocultas separadas, cuja saída é então conectado para produzir a saída final [191]. A estrutura do BiLSTM é ilustrado na Fig. 19. Sua vantagem é que não só considera informações passadas, mas também usa informações futuras do série temporal. No entanto, poucos estudos aplicaram BiLSTM a WS/WP tarefas de previsão. Portanto, a eficácia do BiLSTM não foi amplamente demonstrado.

Biswas e Sinha [192] compararam o desempenho de previsão de Redes LSTM e BiLSTM com diferentes valores de unidades ocultas e valores de época. Os resultados mostraram que uma vez que o BiLSTM utiliza informações de dados passados e futuros na série temporal, tem o potencial para ter um desempenho muito melhor do que LSTM. Liu e Chen [58] desenvolveram um modelo de previsão de conjunto WS usando diferentes modelos BiLSTM como base de preditores e treinou o modelo com dados produzidos por vários métodos de transformação de resolução.

Em [193], AE baseado em BiLSTM foi proposto para extrair importantes recursos não vistos de séries temporais. Em seguida, esses recursos extraídos foram

alimentados em um BiLSTM para obter intervalos de previsão WS. O proposto estratégia alcançou uma melhoria de 39% em termos de largura de cobertura critério comparado com os modelos tradicionais. Em [194], WD foi usado para decompor os dados do WS em várias subséries com diferentes escalas de tempo, seguido por BiLSTM para prever todas as subséries. Essa estratégia gerou previsões mais precisas do que as de quatro modelos de previsão WS tradicionais.

Jahangir et al. [95] utilizaram diferentes modelos BiLSTM para prever WS em diferentes subclusters produzidos por -means clustering.

5.1.3. Memória convolucional de longo prazo

Recentemente, dois métodos de combinação de camada convolucional e LSTM camada foram propostas. O primeiro método aplica uma camada convolucional separada e uma camada LSTM uma após a outra [174,195]. O segundo método, chamado de memória de longo prazo convolucional (ConvLSTM), é uma variante do LSTM usando uma operação de convolução dentro de um LSTM célula [196]. As estruturas de ConvLSTM e LSTM são comparadas em Fig. 20, onde pode ser visto que ConvLSTM substitui a multiplicação de matrizes por uma operação de convolução em cada porta da célula LSTM [196]. Assim, ele captura características espaciais subjacentes de imagens multidimensionais, dados usando operações de convolução.

Wilms et al. [197] utilizaram um ConvLSTM para fazer previsões de WP com base em dados de vento de um local de estudo e pontos vizinhos. o os resultados mostraram que o ConvLSTM pode aprender dependências espaço-temporais dos dados, produzindo assim previsões WP mais precisas do que LSTM. Os desempenhos de previsão do WS de cinco modelos, incluindo ConvLSTM, LSTM, ANN, CNN e SVM foram comparados em [198]. Os resultados indicaram que o ConvLSTM forneceu previsões de WS altamente precisas com menos atrasos e neurônios ocultos, e com menor complexidade computacional.

Sun e Zhao [199] projetaram um modelo híbrido para Previsões WP nas quais ConvLSTM foi empregado para prever todas as subséries geradas pelo VMD, e a tendência na sequência de erros foi modelado pelo LSTM. Em [40], um método baseado em EMD foi empregado para decompor os dados do WS em várias subséries com diferentes frequências. O componente de baixa frequência foi modelado pelo ConvLSTM, e o outros componentes foram previstos pela KELM.

Em [200] e [201], no âmbito da aprendizagem multitarefa, um ConvLSTM foi proposto para prever simultaneamente a produção de energia e carga estrutural de uma turbina eólica. A multitarefa ConvLSTM proposta considerou a correlação entre dois alvos e pôde aprender recursos comuns úteis para prever ambas as saídas de destino usando um único estrutura.

5.2. Rede de unidades recorrentes fechadas e suas variantes

Em 2014, Cho et al. [202] propuseram o GRU, que é uma variante do LSTM e é capaz de aprender dependência longa e curta sem cair em gradientes de fuga e explosão. A estrutura típica de uma GRU unidade é mostrada na Fig. 21. Comparado ao LSTM (Fig. 20), não há célula de memória separada na GRU, e contém apenas duas portas: porta de *reset* e *atualização* do portão. O primeiro ajuda a decidir quanto do anterior estado oculto é considerado no cálculo do candidato do estado atual, enquanto o último combina o portão de *entrada* e o portão de *esquecimento* no LSTM para determinar o grau do estado oculto anterior que será usado para atualizar o estado atual [203]. A estrutura simples do GRU permite que ele seja treinado mais rápido que o LSTM.

Niu et al. [47] projetou um GRU baseado em atenção para fazer à frente das previsões do WP. No modelo proposto, a correlação entre diferentes tarefas de previsão foram consideradas no GRU e uma atenção

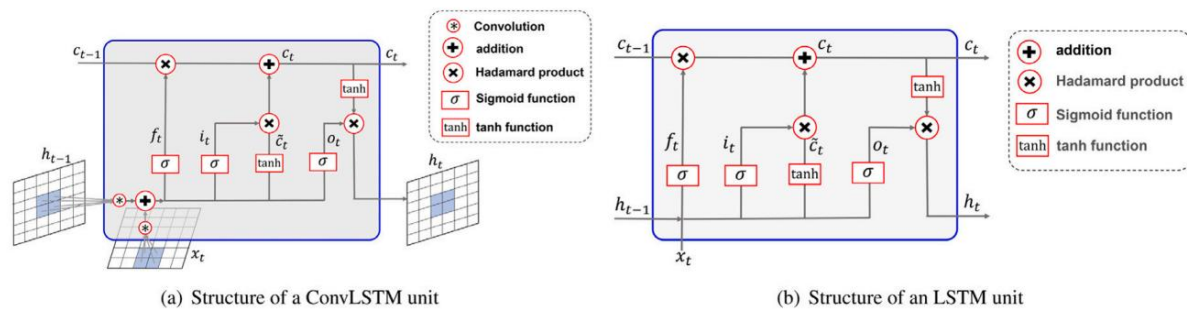


Fig. 20. Estruturas de uma unidade ConvLSTM e uma unidade LSTM.

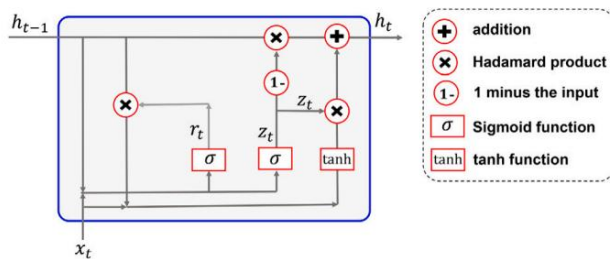


Fig. 21. Estrutura de uma unidade GRU.

mecanismo ajudou a selecionar características importantes. Peng et al. [94] caracterizaram a flutuação em uma série WS sem ruído com um GRU. Em [203], um modelo de previsão WP baseado em GRU forneceu maior precisão de previsão com menor tempo de treinamento do que um modelo baseado em LSTM. Liu et al. [204] decompôs uma série temporal WS em um componente de tendência e vários componentes detalhados. Um CNN-GRU composto por três camadas convolucionais e uma camada GRU foi usado para prever o componente de tendência, enquanto o SVM foi adotado para prever os demais componentes detalhados. O CNN-GRU apresentou desempenho de previsão superior ao de outros sete modelos. Em [175], as características extraídas usando um CCN de todas as subséries de WS foram obtidas por diferentes métodos de decomposição e alimentadas em uma GRU para caracterizar suas dependências temporais em WS futuros. Em [93], todas as subséries decompostas da série WP original foram previstas por uma GRU melhorada cuja função de ativação foi escalada em unidades lineares exponenciais. Ao contrário dos métodos usados em [93], Yu et al. [205] empregaram uma GRU para extrair características profundas de uma subsérie de baixa frequência, enquanto as outras séries foram previstas por diferentes SVMs.

A abordagem GRU também tem sido usada para construir modelos probabilísticos de previsão WS/WP. Afrasiabi [86] construiu um modelo de previsão WS probabilístico híbrido no qual uma rede GRU foi empregada para extrair características temporais das saídas de uma CNN. A função de densidade de probabilidade do WS previsto foi gerada por uma rede de densidade de mistura melhorada. Li et al. [92] projetaram um modelo GRU de conjunto para obter previsões de intervalo WS de curto prazo. No modelo proposto, o VMD foi usado para decompor a série WS original em um modo primário e outros modos simplificados, que foram usados para treinar modelos baseados em GRU para gerar previsões de intervalo e ponto, respectivamente. Em seguida, um modelo de correção de erros baseado em GRU fez as previsões de intervalo final. Em contraste com [92], Wang et al. [206] usaram um modelo baseado em GRU para gerar previsões de intervalo para todos os modos simplificados produzidos pelo VMD, que foram combinados para obter intervalos de previsão de WP. Além disso, a qualidade do intervalo foi melhorada usando um método de otimização adaptativa baseado em intervalo construído [206].

Além do modelo GRU tradicional, variantes como BiGRU e ConvGRU também têm sido usadas na previsão de WS/WP. A estrutura de um BiGRU é a mesma de um BiLSTM. Ding et al. [48] construíram um modelo de previsão baseado em BiGRU para corrigir dados NWP WS, que foram transformados em previsões WP usando uma curva de energia eólica. O proposto

modelo gerou previsões melhores do que SVM e RNAs tradicionais. Xiang et al. [124] decompuseram os dados WS em diferentes subséries com a estratégia SD, então as previsões WS determinísticas e probabilísticas foram obtidas a partir das previsões de todas as subséries geradas por modelos de previsão baseados em BiGRU. Em [179], um ConvGRU e um 3D-CNN foram combinados para fazer previsões de WS espaço-temporais dentro de uma previsão de codificação. estrutura.

5.3. Outros modelos baseados em aprendizado profundo

Além de modelos DNN populares, como GRU, LSTM e suas variantes, algumas outras técnicas profundas também foram empregadas para previsão de WS/WP.

Zhang et al. [207] projetou uma rede de densidade de mistura profunda aprimorada (IDMDN) para previsão probabilística WP regional de curto prazo. Em um IDMDN, um kernel beta é selecionado como o componente da mistura para evitar o problema de vazamento de densidade. Além disso, para superar o comportamento instável e resolver problemas de falta de valor da função do kernel beta, uma função de ativação e procedimentos de treinamento também foram desenvolvidos e definidos.

Motivados pela aproximação de primeira ordem localizada de convoluções de grafos espectrais, Khodayar e Wang [208] desenvolveram uma arquitetura de aprendizado profundo convolucional de grafos escalável baseada na teoria dos grafos, CNN e teoria dos conjuntos brutos para extrair características espaciais e temporais de dados WS.

Ao contrário das redes neurais tradicionais, as redes Bayesianas podem descrever a incerteza do modelo. Wen et al. [209] desenvolveram um modelo de aprendizado profundo Bayesiano para fazer previsões probabilísticas de WP, com o dropout de Monte Carlo utilizado para aproximar a distribuição posterior dos parâmetros do modelo. Em [179], um dropout de Bernoulli foi usado na implementação de um modelo de aprendizado profundo baseado em Bayesiano variacional para fazer previsões probabilísticas espaço-temporais de WS.

Hu et al. [210] estendeu a estrutura de aprendizado profundo para uma rede básica de estado de eco (ESN), o que resultou em um ESN profundo que compreendia uma hierarquia empilhada de reservatórios. As vantagens do ESN profundo proposto incluem maior capacidade de ajustar séries temporais não lineares e uma estrutura de aprendizado profundo mais eficiente, o que tornou o ESN profundo mais preciso para previsão de WP em comparação com o ESN original.

Liu et al. [211] desenvolveram uma abordagem de aprendizado por reforço profundo para previsão de WS de curto prazo. Especificamente, um método Q-learning foi empregado para integrar três redes neurais: LSTM, DBN e ESN. Essa integração trouxe melhor desempenho na previsão de WS de curto prazo do que um método tradicional de ensemble baseado em otimização.

Zhang et al. [87] usaram sequência a sequência (Seq2Seq), uma técnica de aprendizado profundo popular em tradução automática e diálogo homem-máquina, para mapear uma série temporal de entrada para uma série temporal WP de saída usando valores de tag e atenção. Os resultados mostraram que o método baseado em Seq2Seq teve um desempenho melhor do que as técnicas de DBN e floresta aleatória (RF) na previsão de WP com antecedência.

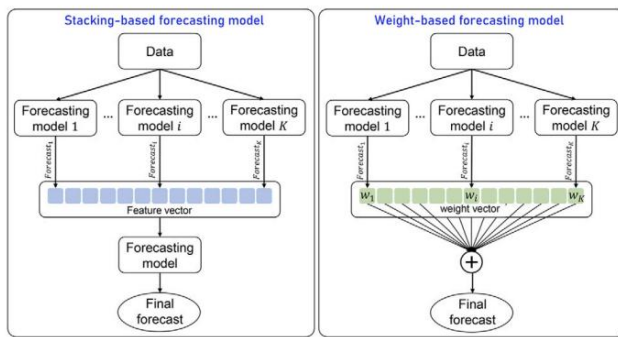


Fig. 22. Estrutura de modelos baseados em empilhamento e modelos baseados em peso.

Tabela 4

Redes neurais profundas híbridas na literatura.

Tipo híbrido	Artigo	Previsão objetivo	Modelos híbridos
Baseado em empilhamento modelo de previsão	[57]	WS	LSTM + SVM
	[55,56]	WS	empilhado LSTM
	[63]	WS	LSTM + LEME
	[59]	WP	AE + DBN
	[205]	WS	RNN + SVM
	[169]	WS	LSTM + SVM
	[212]	WS	GRU + SVM
	[19]	WP [53]	DBN + RF
	[54]	WS	DBN + LGBM
	[213]	WS	SAE + ELM
	[204]	WS	CNN + LSTM
	[214]	WP	CNN + LSTM
	[114]	WS	CNN + LSTM
	[179]	WS	CNN + LSTM
	[213]	WS [98]	CNN + BiLSTM
		WS	SDAE + LSTM
	[44]	WS	MLP + LSTM + ARIMA
	[215]	WP	GRU + RKELM
	[216]	WS	vários NCULSTMs
	[109]	WS [58]	LSTM + Elman
Com base no peso modelo de previsão	[217]	WS	SSAE-BiLSTM + BiLSTM
	[204]	WS	GRU + RBFNN
	[107]	WP	CNN-GRU + SVM
	[122]	WS [39]	DBN + LSSVM
	[40]	WS	LSTM + DBN
	[124]	WS	LSTM + ELM
	[18]	WP	ConvLSTM + KELM
	[178]	WP	BiGRUs
	[218]	WS	SDAEs
	[219]	WS	CNNs
	[220]	WS	LSTM + ARIMA
	[221]	WS	ARIMA + BPNN + DBN + ELM
	[222]	WS	+ Elman + GRNN + LSTM

5.4. Redes neurais profundas híbridas

Muitos modelos híbridos têm sido propostos para explorar as vantagens de diferentes modelos. Os modelos de previsão híbridos são divididos em dois categorias: modelos baseados em empilhamento e modelos baseados em peso [52], cujas estruturas são comparadas na Fig. 22. Alguns modelos híbridos baseados em DNN em a literatura é apresentada na Tabela 4.

Modelos baseados em empilhamento, como mostrado na Fig. 22, geralmente tratam as previsões de um ou mais modelos de base como recursos a serem combinados com outro modelo de previsão. Em [57], seis LSTMs diversos foram propostos para explorar e explorar as informações implícitas nos dados do WS e, em seguida, o

previsões de diferentes LSTMs foram combinadas pelo SVM, cujos parâmetros foram ajustados pelo algoritmo de otimização extrema (EO). Em [63], vários Os modelos LSTM e histerético ELM (HELM) foram empregados para gerar previsões básicas que foram combinadas pelo LSTM com camadas ocultas e neurônios otimizados por evolução diferencial (DE). Em [59], nove EAs foram empregados como base-regressores, enquanto DBN atuou como meta-regressor para gerar as previsões finais do WP. Em [205], as saídas de RNN, GRU, LSTM foram alimentados no SVM para prever a sub-série de aproximação obtidos por WD para dados de WS. Semelhante a [205], os autores de [169] também empregou WD para decompor a série temporal WS em diferentes sub-séries, em que as características extraídas pelo DBN foram alimentadas no LGBM e RF. Em [212], o SAE foi aplicado para extrair recursos e o ELM foi usado para prever o futuro WS. Em [204], a CNN seguida pela GRU, referida como CNN-GRU, foi empregado para prever o componente de tendência dos dados do WS. Ao contrário da abordagem usada em [204], a CNN foi usada para extrair recursos, que foram então alimentados no SVM para modelar as sub-séries geradas por decompondo os dados do WS com EMD. Para melhorar a previsão do WS precisão, vários LSTMs foram empilhados para formar modelos híbridos para fazer uso pleno de informações espaço-temporais [55] e dados meteorológicos múltiplos [56], respectivamente. Em [214], um ConvLSTM dilatado empilhado foi desenvolvido para prever a resposta de turbinas eólicas em um campo eólico complexo. Em [19,53] e [54], as feições espaço-temporais foram extraídas por uma estrutura profunda híbrida integrando uma CNN e uma LSTM. De forma similar, CNN e BiLSTM também foram hibridizados para extrair recursos para gerar as previsões finais do WS. Em [179], ConvGRU e 3D-CNN foram combinados dentro de uma estrutura de previsão de codificação para produzir as previsões espaço-temporais. Em [98], a informação intrínseca contidos nos dados de vento originais foram extraídos pelo SDAE, e o LSTM foi utilizado para produzir as previsões finais do WS.

Modelos baseados em peso, conforme mostrado na Fig. 22, atribuem pesos diferentes para combinar previsões de modelos base. Os tipos de pesos usados mais divide os modelos baseados em peso em modelos baseados em peso fixo e modelos dinâmicos baseados em peso.

Para modelos baseados em peso fixo, cada peso recebe um valor de 1 ou o inverso do número de modelos básicos de previsão. Fazendo assim, as previsões finais podem ser calculadas através de média simples ou soma. Em [215], EWT foi usado para decompor a série temporal WP em diferentes sub-séries, em que a sub-série com baixa entropia amostral foram previstos com GRU, enquanto as sub-séries com alta entropia amostral foram previstas pelo kernel reduzido ELM (RKELM) otimizado pelo algoritmo de molde de lodo (SMA). As previsões finais do WP foram obtidas acumulando as saídas de todos os modelos de previsão. Em [216], um romance LSTM chamado New Cell Update LSTM (NCULSTM) foi proposto e treinados com as sub-séries geradas pelo EWT, cujas previsões foram agregados para obter as previsões finais do WS. Em [109], LSTM e Elman foram usados para prever as frequências de baixa e alta frequência sub-série, respectivamente. Em [204], CNN-GRU e SVM foram combinados para prever o componente de tendência e detalhar os componentes dos dados do WS, enquanto em [107], LSSVM foi usado para prever o componente de tendência de dados WP e DBN para os componentes de detalhes correspondentes. Em [122], O LSTM foi adotado para prever a sub-série WS de baixa frequência, enquanto DBN de três camadas com hiperparâmetros otimizados com gafanhoto algoritmo de otimização (GOA) foi empregado para prever a sub-série de alta frequência. Da mesma forma, o LSTM também foi usado para modelar o sub-série de baixa frequência e ELM para sub-série de alta frequência em [39]. Em [40], dois preditores, ConvLSTM e KELM, foram combinados para previsão de sub-séries, e os parâmetros ótimos em KELM foram selecionados por um novo algoritmo (MHOGWO) que combinou Harris Hawks otimização e o otimizador de lobo cinzento. Um cluster de BiGRUs cujo parâmetros foram otimizados pela otimização de enxame de frango (CSO) foram combinados para prever as sub-séries WS geradas por um método SD [124]. Em [178], as sub-séries WP obtidas por WD foram previstas por um conjunto de CNNs. Em [218], EMD decompôs os dados de vento originais em várias sub-séries com diferentes frequências e um residual; a sub-séries de alta frequência foram previstas com LSTM e o restante foram modelados com ARIMA.

Para modelos baseados em peso fixo mais complexos, abordagens de otimização inteligente foram adotadas para selecionar pesos ideais. Um ensemble de MLP, LSTM e ARIMA foi construído em [44] para previsão de WS, com o peso ótimo de cada modelo obtido através do algoritmo de otimização de enxame de partículas multiobjetivo (MOPSO).

Em [58], um ensemble de dois previsores de base foi construído para fazer pleno uso das informações em dados de vento multi-resolução, em que o BiLSTM foi selecionado para dados de vento de baixa resolução, enquanto o BiLSTM com as características extraídas pelo SSAE foi construído para dados de vento de alta resolução. dados de resolução. Os pesos dos dois previsores de base propostos foram ajustados pela otimização multiobjetivo multiuniverso (MOMVO). Em [219], cinco modelos ótimos de previsão de WS foram selecionados a partir de sete candidatos, ARIMA, BPNN, DBN, ELM, Elman, GRNN e LSTM, e os pesos dos selecionados foram otimizados pelo algoritmo de libélula multiobjetivo modificado (MMODA). Em [222], a GRU foi empregada para prever cada subsérie gerada pelo EWT, enquanto a rede neural de pico convolucional (CSNN) foi usada para modelar a série de erros de previsão correspondente. As previsões de cada subsérie, que podem ser vistas como a soma das previsões de GRU e CSNN, foram combinadas para produzir as previsões finais de WS, e os pesos ótimos foram determinados pelo otimizador de lobo cinzento (GWO).

Em modelos dinâmicos baseados em peso, o peso de cada predictor de base varia com as amostras. Em [217], a subsérie gerada pelo CEEMDAN foi alimentada em cinco GRUs, seguidas pelo RBFNN otimizado pelo algoritmo Bat melhorado (IBA) para combinar as previsões de diferentes componentes para produzir as previsões finais do WS. Em [18], uma série de SDAEs com diversos parâmetros de modelo e recursos de entrada foram combinados para formar um conjunto para previsão WP de vários parques eólicos, e a regressão linear multi-resposta (MLR) foi adotada como meta-aprendiz para treinar o conjunto SDAE.

Em [220], o WS foi previsto primeiro com quatro preditores de base: BPNN em cascata, ESN, rede de entradas exógenas (NARX) e LSSVM, e os resultados dos previsores de base foram combinados usando ESN otimizado por otimização de enxame de partículas de comportamento quântico (QPSO). Da mesma forma, em [221], três RNAs, GRNN, Elman e cascata BPNN (CBP) foram empregadas para prever WS inicialmente, cujos resultados foram combinados com ESN otimizados através do MOPSO para gerar as previsões finais de WP.

Modelos híbridos, que levam em conta as características de diferentes modelos, geralmente apresentam desempenho superior na descrição de relacionamentos complexos em comparação com modelos únicos. Assim, não é surpresa que diferentes tipos de DNNs híbridas tenham sido amplamente utilizados para previsão de WS e WP na literatura.

5.5. Comparação de diferentes métodos de aprendizado de relacionamento baseados em DNN

Diferentemente das RNAs tradicionais, as RNNs podem minerar as informações temporais nos dados sequenciais, o que torna esses modelos mais eficazes para lidar com as tarefas relacionadas ao processamento de sequências, principalmente na previsão de séries temporais. Como resultado, diferentes tipos de RNNs são amplamente utilizados para caracterizar a flutuação complexa de WS/WP na previsão de WS/WP. Este artigo compara principalmente quatro RNNs populares: LSTM, GRU, BiLSTM e BiGRU.

A maior vantagem do LSTM sobre o RNN tradicional está em sua capacidade de aprender dependências de longo prazo e evitar o desaparecimento de gradientes com três tipos diferentes de portas [223]. Os pares GRU esquecem a porta e a porta de entrada [224]. Em primeiro lugar, uma unidade LSTM produz três valores enquanto GRU tem apenas dois valores para saída. Embora o LSTM precise de mais parâmetros de treinamento, use mais memória e tenha velocidade de treinamento mais lenta do que o GRU, ele é mais preciso para conjuntos de dados com sequências mais longas [224]. Resumindo, o LSTM é mais adequado para grandes sequências ou onde é necessária alta precisão, enquanto o GRU é favorável quando menos memória, mas é necessária uma operação mais rápida.

Modelos bidirecionais usam informações passadas e futuras para fazer previsões [225]. Recursos mais úteis geralmente permitem BiLSTM e BiGRU

para ter um desempenho melhor do que o normal LSTM e GRU, respectivamente. No entanto, células duplas LSTM ou GRU tornam o BiLSTM e o BiGRU caros.

Diferentes estratégias híbridas levam a diferentes custos computacionais. Os custos computacionais são altos para modelos híbridos baseados em empilhamento, pois diferentes componentes são conectados em série, o que significa que as informações passam por cada componente um a um. No entanto, em modelos híbridos baseados em peso, todos os modelos básicos são conectados em paralelo. Modelos com modo de computação paralela consomem menos tempo do que modelos com modo de computação serial. Portanto, modelos baseados em peso podem ser treinados mais rapidamente do que modelos baseados em empilhamento.

Como os modelos híbridos podem aproveitar ao máximo as vantagens de diferentes modelos, eles geralmente têm um desempenho melhor do que os modelos únicos correspondentes. No entanto, os custos computacionais dos modelos híbridos baseados em empilhamento são maiores do que os modelos únicos, e os modelos híbridos baseados em peso são comparáveis com os modelos únicos em termos de custos computacionais quando são treinados em paralelo. Em aplicações reais, tanto os custos computacionais quanto a precisão da modelagem devem ser considerados simultaneamente para selecionar modelos adequados para caracterizar as relações complexas em os dados.

6. Otimização da configuração do modelo

O desempenho de diferentes modelos de previsão é fortemente influenciado pelas configurações do modelo, como arquitetura do modelo e hiperparâmetros. Para maximizar seu desempenho, muitos algoritmos de otimização inteligentes foram desenvolvidos e adotados para otimizar vários aspectos dos modelos de previsão ao longo das últimas décadas [46]. Alguns estudos relacionados foram citados na Seção 5.4.

Mais estudos sobre a determinação de configuração de modelo ideal são revisados a seguir, e a Tabela 5 resume alguns otimizadores usados nos estudos anteriores.

Embora os modelos de previsão baseados em kernel, como SVM, LSSVM e KELM, tenham sido amplamente empregados para previsão de WS/WP, pois apresentam desempenho superior em lidar com problemas de regressão não linear, é uma tarefa desafiadora selecionar os parâmetros ideais do modelo, incluindo o tipo de função de kernel, os parâmetros de kernel correspondentes e o coeficiente de regularização. Em [141] e [149], GA e PSO foram adotados para sintonizar SVM e SVM reduzido, respectivamente.

Em [226], o -SVM baseado em CS foi proposto para melhorar a precisão da previsão de WS. Em [61], os parâmetros ótimos em SVM foram obtidos por um IASA, enquanto EO foi adotado em [57]. Em [38], o otimizador híbrido PSO-GSA foi usado para escolher parâmetros ótimos de LSSVM com velocidade de convergência mais rápida que PSO. Da mesma forma, o PSO baseado em recozimento simulado (PSOSA) também foi usado para ajustar o LSSVM [227]. Em [228], uma otimização de evolução diferencial híbrida aprimorada-Harris Hawks (DE-HHO) foi proposta para ajustar KELM para previsão de WS.

Quando se trata de modelos de previsão baseados em RNAs, incluindo as tradicionais RNAs rasas e as populares DNNs, é importante determinar as configurações apropriadas do modelo, por exemplo, pesos, limites, número de camadas ocultas, número de neurônios em cada camada, lote tamanho, taxa de aprendizado e tipo de otimizador [46,63]. Em [229], a potência de saída de turbinas eólicas foi estimada por ELM, cujos pesos e vieses foram determinados por otimização de colônia de formigas (ACO). O algoritmo cruzado (CSA) foi aplicado para ajustar os pesos de entrada e os vieses da camada oculta no ELM robusto de outlier sequencial online (ORELM) e ELM comum em [107] e [235], respectivamente. Os pesos de saída do ELM com um nó oculto foram otimizados com um método de evolução diferencial autoadaptativa (SADE) para fazer previsões WP probabilísticas [238].

Em [232], um otimizador de lobo cinza híbrido aprimorado e algoritmo de seno cosseno (IHGWOSCA) foram usados para otimização de ELM. Em [60], BPNN otimizado por GA, WNN otimizado por CS e GRNN otimizado por CG-BA foram usados para fazer previsões de WS de curto prazo para fins de comparação.

Em [230] e [231], o BPNN otimizado por AG foi combinado com técnicas de decomposição para prever o futuro WS. Em [234], um novo otimizador combinou otimização adaptativa conduzida pelo vento e recozimento simulado modificado (AWDO-MSA) para otimizar pesos iniciais e limites de

Tabela 5
Otimizadores para configuração de modelos.

Tipo de modelo	Modelos	Parâmetros	Otimizador
Modelos baseados em kernel	SVM, LSSVM, KELM	Coeficiente de regularização dos parâmetros do kernel etc.	GA [141], PSO [149], CS [226], ASA [61], IASA [61], PSOGSA, [38], EO [57], PSOSA [227], DE-HHO [228], SMA [215], MHOGWO [40],
	SVM reduzido, -SVM RKELM		
Modelos baseados em ANN	ELM, BPNN, GRNN	Pesos, vieses, números de taxa de aprendizado de camadas ocultas, números de neurônios ocultos, tamanho do lote, raio espectral do tamanho do pool de armazenamento, etc.	ACO [229], CG-BA [60], GA [60,62,136,230,231], CS [60], SADE [232], IHGWOSCA [232], FA [233], MOSCA [64], IBA [217], DE [63], GNDO [46], GOA [122], AWDO-MSA [234], CSA [107.235], CSO [124], QPSO [220], MOSCOU [64], MOCS [65], MOGWO [66], MOPSO [221]
	WNN, ORELM, Elman RBFNN, MLP, LSTM DBN, BiGRU, BiLSTM, SDAE, ESN		
Modelos de previsão híbridos baseados em peso	/	Pesos combinados, etc.	GWO [43.222.236], MOCS [64], MOBA [237], MOMVO [58], MOPSO [44] MMODA [219]

um modelo Elman. Foi utilizado um MLP com o algoritmo Firefly (FA) para previsão de WS de longo prazo em [233]. Em [62], o LSTM foi ajustado por GA, enquanto em [63], o algoritmo DE foi usado para LSTM para balancear sua complexidade. Em [46], uma nova versão do GNDO foi empregada para ajustar a arquitetura e os hiperparâmetros do BiLSTM. Em [136], GA foi usado para selecionar a configuração ideal do SDAE para extrair os recursos importantes dos dados do WS. Comparado com o objetivo único métodos de otimização, otimizadores multi-objetivos ajudam a alcançar tanto alta precisão e forte estabilidade na previsão simultaneamente [64]. Com a esperança de gerar previsões WS/WP precisas e estáveis, MOSCA [64], MOCS [65] e MOGWO [66] foram empregados para otimizar os modelos WNN, Elman e ELM, respectivamente.

Em modelos de previsão híbridos baseados em peso, o peso de cada base Previsor desempenha um papel importante na precisão final da previsão. Em [236], o LSTM foi usado para prever cada função de modo intrínseco gerado pelo ICEEMDAN, e as previsões para todos os componentes foram combinado com pesos determinados por GWO. Da mesma forma, em [43], GWO também foi adotado para escolher os pesos ótimos para combinação de BPNN, Elman, WNN e GRNN para previsão de WS. Em [64], um conjunto de filtros Elman e Volterra foi construído para fazer WS previsões com os pesos decididos pelo MOCS e teoria das restrições não-negativas. Em [237], os pesos ótimos de quatro RNAs foram ajustados pelo algoritmo de morcego multiobjetivo (MOBA). Em [58], os pesos ótimos de dois previsores de base BiLSTM foram obtidos por otimização multi-objetivo multi-universo (MOMVO). Em [239], um novo algoritmo de otimização BWCA combinando o algoritmo do ciclo da água e a família Broyden algoritmo, foi proposto para ajustar os pesos de cinco redes para melhorar a precisão e a estabilidade da previsão do WS.

Revisão de estudos anteriores para determinação do modelo ideal configurações revela que otimizadores inteligentes têm sido amplamente utilizados em muitos aspectos da previsão de WS/WP. No entanto, um pequeno número de estudos têm usado algoritmos de otimização inteligente para obter a configurações ideais para DNNs, que, na maioria dos casos, são decididas por tentativa e erro manualmente [46.217]. A razão para o uso limitado de algoritmos de otimização inteligente para decidir as configurações ideais para DNNs pode levar muito tempo para ajustar o modelo, pois DNNs com vários parâmetros de modelo são altamente complexos.

7. Estudos de caso numéricos para diferentes métodos

Nesta seção, alguns modelos populares de previsão de WS/WP são comparados usando dois conjuntos de dados do mundo real.

7.1. Descrição de dados

Os dados WS e WP usados aqui estão disponíveis em sites abertos (<https://maps.nrel.gov/wind-prospectors/> e <https://data.open-power-system-data.org/>, respectivamente). O conjunto de dados WS foi coletado de 1º de janeiro de 2012 a 31 de dezembro de 2012 com uma resolução de tempo de 30 min. Os primeiros 70% dos pontos de dados foram usados como treinamento amostras, e os pontos de dados restantes foram usados como dados de teste. O WP O conjunto de dados foi coletado de 1º de janeiro de 2016 a 31 de dezembro de 2016 com resolução temporal de 1 h. O conjunto de treinamento e o conjunto de teste também foram dividido de acordo com a razão de 7 ÷ 3.

Tabela 6
Equações para cinco indicadores de erro usados no artigo.

Indicador de erro	Cálculo
MUITO DE	$\frac{1}{n} \sum_{i=1}^n \hat{y}_i - y_i $
RMSE	$\sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}$
MAPA	$\frac{1}{n} \sum_{i=1}^n \left \frac{\hat{y}_i - y_i}{y_i} \right \times 100\%$
SMAP	$\frac{1}{n} \sum_{i=1}^n \frac{ \hat{y}_i - y_i }{\frac{1}{n} \sum_{i=1}^n \hat{y}_i + \frac{1}{n} \sum_{i=1}^n y_i } \times 100\%$
2	$\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2$

Nota: \bar{y} é a média das amostras de teste com o \hat{y} , e e são as valor real da contagem e o valor estimado correspondente, respectivamente.

7.2. Configuração experimental

Os modelos comparados nos estudos de caso podem ser divididos em três grupos: modelos tradicionais de benchmark, modelos de profundidade única e modelos de profundidade híbrida. Três modelos tradicionais de benchmark foram selecionados, ou seja, modelo de persistência, regressão linear (LR) e BPNN. O solteiro modelos profundos usados aqui foram LSTM, GRU, BiLSTM, BiGRU, 1D-CNN e 2D-CNN. Para comparar os modelos híbridos, foram selecionados oito modelos de previsão baseados em empilhamento. Eles eram 1D-CNN-LSTM, 1D-CNN-GRU, 1D-CNN-BiLSTM, 1D-CNN-BiGRU, 2D-CNN-LSTM, 2D-CNN-GRU, 2D CNN-BiLSTM e 2D-CNN-BiGRU. Nesses modelos híbridos, a CNN foi usado para extrair recursos, que foram então alimentados nos outros modelos profundos para fazer previsões de WS/WP. Todos os modelos baseados em aprendizado profundo foram implementado em Keras [240] com função de perda "mse" e otimizador "adam", e suas configurações foram determinadas por trilha e erro.

Para medir o desempenho de previsão de diferentes modelos, cinco indicadores de erro de previsão foram usados: erro absoluto médio (MAE), erro quadrático médio (RMSE), porcentagem absoluta média erro (MAPE), erro percentual absoluto médio simétrico (SMAPE) e coeficiente de determinação (R^2). As equações para esses indicadores são dado na Tabela 6.

7.3. Previsão de resultados e discussões

Os resultados de previsão WS/WP de diferentes modelos aplicados a dois conjuntos de dados são apresentados na Tabela 7. Pode-se ver que dois modelos de benchmark, o modelo de persistência e LR, têm desempenho pior do que BPNN e outras DNNs. Esse fenômeno mostra que os modelos não lineares caracterizam as flutuações de WS/WP com mais precisão e são mais adequados do que os modelos lineares para previsão WS/WP.

Em termos da maioria dos indicadores de erro, a CNN alcança o melhor desempenho de previsão entre DNNs únicos, incluindo LSTM, GRU, BiL STM, BiGRU, 1D-CNN e 2D-CNN. Para previsão de WS, o MAE, RMSE, MAPE, SMAPE e 2 de 1D-CNN são 0,4794 m/s, 0,7308 m/s, 9,8476%, 4,3109% e 0,9630, respectivamente, que são obviamente

Tabela 7
Previsão de resultados de diferentes modelos para dois conjuntos de dados.

Método	Previsão de WS					Previsão de WP				
	MUITO DE	RMSE	MAPA	SRMSE	2	MUITO DE	RMSE	MAPA	SRMSE	2
Persistência	0,5405	0,8048	10,0464	4,8273	0,9552	200.0679	345.4666	6,6693	3,2243	0,9747
LR	0,5133	0,7706	10.1503	4,5801	0,9589	198.9332	343.1313	6,8459	3,2858	0,9750
BPNN	0,5103	0,7693	10.1719	4,5465	0,9591	178.7424	320.8060	6,0750	2,9679	0,9782
LSTM	0,5163	0,7744	10.5124	4,6755	0,9585	176.5142	326.5269	6,0719	3,0210	0,9774
GRU	0,5108	0,7692	10.1235	4,5716	0,9552	176.8537	326.9409	6.1089	3,0671	0,9773
BiLSTM	0,5184	0,7755	10.0760	4,6077	0,9584	175.8823	323.4891	5,9888	2,9571	0,9778
BiGRU	0,5119	0,7693	10.1230	4,5729	0,9590	176.9383	326.7199	6,1000	3,0513	0,9774
1D-CNN	0,4794	0,7308	9,8476	4,3109	0,9630	171.6678	313.3751	5,8942	2,8386	0,9792
2D-CNN	0,4805	0,7315	9,8101	4,3166	0,9630	173.6291	311.2748	6,0395	2,9031	0,9795
1D-CNN-LSTM	0,4738	0,7289	9,9316	4,2790	0,9632	171.7344	321.6857	6,0238	2,8496	0,9781
1D-CNN-GRU	0,4743	0,7315	9,7764	4,2710	0,9630	170,9881	318.1121	5,9359	2,8245	0,9785
1D-CNN-BiLSTM	0,4751	0,7302	9,5507	4,2559	0,9631	173,1694	320.3462	6,0803	2,8813	0,9782
1D-CNN-BiGRU	0,4733	0,7305	9,6160	4,2407	0,9631	171.3352	317.5747	5,9452	2,8248	0,9786
2D-CNN-LSTM	0,4767	0,7352	9,7090	4,2620	0,9626	175.6808	310.4976	6,2563	2,9699	0,9796
2D-CNN-GRU	0,4798	0,7357	9,7472	4,3138	0,9625	175.4367	313.2930	6,4608	2,9973	0,9792
2D-CNN-BiLSTM	0,4816	0,7426	10.2982	4,3895	0,9618	175.5819	310.6278	6,2430	2,9751	0,9795
2D-CNN-BiGRU	0,4783	0,7331	9,8070	4,2981	0,9628	174.4334	309.8374	6.1851	2,9633	0,9796

melhor do que os de LSTM, GRU, BiLSTM e BiGRU. Para previsão WP, o 1D-CNN também alcança um excelente desempenho com MAE, MAPE, SMAPE e 2 sendo 171,6678 MW, 5,8942%, 2,8386%, e 0,9792, respectivamente. Além disso, pode-se observar na Tabela 7 que o 1D-CNN e o 2D-CNN têm desempenho comparável, e ambos superam as outras DNNs simples, o que demonstra que o recursos extraídos pela CNN são adequados para explicar com precisão o flutuações de WS e WP.

Também é evidente na Tabela 7 que o desempenho de previsão de Os modelos híbridos baseados em CNN são superiores aos DNNs únicos. Por exemplo, 1D CNN-LSTM e 2D-CNN-LSTM são caracterizados por valores de MAE em Previsões do WS que são reduzidas em 8,2316% e 7,6700%, respectivamente, comparado ao LSTM. Em termos de MAE, a melhoria de 1D-CNN BiGRU e 2D-CNN-BiGRU atinge 7,5405% e 6,5638% em comparação com o BiGRU. Resultados semelhantes também podem ser obtidos para a previsão de WP. A razão para esses fenômenos pode ser que as características extraídas por CNN ajuda a identificar as dependências temporais, melhorar o desempenho de previsão de modelos únicos. No mesmo vez, esses resultados mais uma vez confirmam a validade das características capturado pela CNN.

Entre os diferentes modelos híbridos, pode-se observar que os modelos baseados em 1D-CNN têm um desempenho melhor do que os modelos baseados em 2D-CNN na maioria dos casos. Além disso, 1D-CNN tem um desempenho um pouco melhor do que 2D-CNN em termos de indicadores de erro. Esses resultados mostram que 1D-CNN pode ser melhor do que 2D-CNN na extração de recursos úteis de séries temporais unidimensionais.

Em resumo, a maioria dos modelos de previsão baseados em deep learning são superiores aos métodos tradicionais de previsão WS/WP. Além disso, baseado na CNN modelos híbridos têm melhor desempenho de previsão do que modelos únicos sem CNN. 1D-CNN pode ser uma boa escolha para previsão com dados WS/WP unidimensionais.

8. Desafios e tendências futuras na previsão de WS/WP

Nesta seção, os desafios na previsão de WS/WP são discutidos e são apresentadas as tendências futuras para melhorar a precisão da previsão de WS/WP.

8.1. Desafios na previsão de WS/ WP

Existem três desafios principais na previsão de WS/WP.
(1) Lidar com incertezas complexas em dados de vento.
Incertezas complexas reduzem a previsibilidade dos dados observados, evitando assim que os modelos de regressão aprendam a relação real entre recursos de entrada e previsões WS/WP. Muitos métodos podem ser usado para pré-processar tais incertezas e sua eficácia tem sido comprovada em muitos casos.
Os métodos de pré-processamento geralmente aumentam a precisão da previsão, mas para diferentes graus. Não está claro que um pré-processamento específico

método sempre supera os outros. A razão pode ser que os laços incertos são muito complexos e variam de acordo com o conjunto de dados.
determinado método de pré-processamento pode ser adequado apenas para lidar com uma incerteza específica. Portanto, os métodos atuais de pré-processamento de dados são limitados pela complexidade das incertezas e pela inconsistência entre as incertezas reais e as incertezas que são aplicáveis para modelos de pré-processamento.

(2) Encontrar e extrair recursos úteis de forma eficiente.

Um fator importante que afeta a precisão da previsão de WS/WP é a recursos de entrada. É bem conhecido que as flutuações no WS/WP são afetadas por vários fatores que contribuem para as variações de WS/WP e resultam em previsões imprecisas. Além disso, resultados insatisfatórios das previsões de WS/WP da disponibilidade limitada de dados, tornando essencial tentar coletar mais dados.

Em casos de dados limitados, a mineração de um número maior de dados mais eficazes recursos não apenas ajudam a explicar as flutuações de WS/WP com mais precisão mas também melhora as eficiências de treinamento e previsão de regressão modelos. As maiores vantagens dos métodos clássicos de seleção de recursos são sua fácil implementação e baixa complexidade computacional. O máximo de eles podem ajudar a selecionar recursos úteis de candidatos a recursos. No entanto, sua desvantagem é que eles não podem lidar com recursos complexos profundos em dados de vento que possuem características altamente não lineares. Esta desvantagem está ausente dos métodos de extração de recursos baseados em DNN que, no por outro lado, possuem maior complexidade computacional e apresentam dificuldades na determinação da estrutura ótima da DNN. Portanto, descobrir como obter recursos úteis de forma eficaz é a chave para gerando previsões precisas de WS/WP em um tempo razoável.

(3) Caracterizando as relações complexas entre entrada recursos e dados WS/WP direcionados de forma automática e robusta.

Embora as DNNs teoricamente sejam poderosas para modelar relações não lineares, seu desempenho depende muito do modelo configuração, que geralmente é determinada por tentativa e erro de acordo com os dados futuros de WS/WP. Além da busca demorada para modelos ótimos, não é razoável na prática usar DNNs com dados WS/WP futuros desconhecidos para ajudar a determinar o modelo ideal estrutura e parâmetros. Por isso, torna-se cada vez mais importante saber como determinar o modelo ótimo automaticamente para um determinado conjunto de dados.

Na prática, os relacionamentos de dados mudam com a tarefa de previsão e dados observados, e a capacidade de previsão superior de DNNs não pode sempre ser garantida quando existem relações não lineares complexas. Nesses casos, saber como projetar um modelo adequado baseado em DNN que melhor descreve relações não lineares complexas é essencial para Previsão de WS/WP.

8.2. Tendências futuras na previsão de WS/WP

As seguintes estratégias podem ser consideradas para superar os desafios descritos acima e aumentar a precisão da previsão de WS/WP.

(1) Desenvolvimento de novos métodos de pré-processamento de dados que levem em conta as características de incertezas nos dados de vento.

Ao contrário das incertezas em outros conjuntos de dados, as incertezas nos dados de vento não foram suficientemente analisadas. Portanto, as suposições subjacentes de incerteza que são usadas em métodos de pré-processamento de dados podem não ser válidas quando aplicadas a dados de vento, o que significa que análises baseadas em métodos de pré-processamento de dados existentes podem não ser adequadas para dados de vento, resultando em conclusões tendenciosas. Portanto, há uma necessidade urgente de desenvolver métodos de pré-processamento de dados apropriados para lidar com incertezas complexas em dados de vento. Diferentes tipos de abordagens de pré-processamento de dados são necessários para lidar com esse problema.

(2) Aumentar o número de recursos candidatos que contribuem para flutuações no WS/WP. Na prática, muitos fatores podem causar flutuações no WS/WP. Portanto, usar mais fatores, como fatores topográficos e meteorológicos, será útil para descrever com mais precisão as variações de WS e WP em um local de estudo. Esses dados podem ser coletados instalando sensores adequados no local.

(3) Projetar métodos apropriados de extração de recursos para extrair recursos mais úteis de recursos limitados. As características de métodos clássicos de extração de recursos e modelos baseados em DNN dão à sua combinação um grande potencial para explorar recursos mais úteis. Por exemplo, os métodos clássicos selecionam recursos rasos de um reservatório de recursos, enquanto os modelos baseados em DNN ajudam a gerar expressões de recursos profundos para recursos rasos. Diferentes métodos podem ser usados para extrair diferentes tipos de recursos, que podem se complementar para extrair recursos mais úteis.

(4) Melhorar a capacidade de ajuste não linear dos modelos de previsão. Geralmente, a variabilidade climática e o terreno complexo dificultam a caracterização das flutuações no WS/WP. Embora os modelos de previsão baseados em DNN tenham uma capacidade de ajuste não linear superior, seu desempenho é restrito em aplicações reais devido à complexidade dos dados de treinamento e sua capacidade de aprendizado limitada. Portanto, é útil combinar diferentes modelos de previsão para melhorar a capacidade de ajuste não linear. Por exemplo, diferentes modelos de previsão baseados em DNN e tradicionais podem ser usados para caracterizar a natureza flutuante de WS/WP de forma independente. Então, uma estratégia combinada de forma otimizada pode ser desenvolvida para aproveitar as diferentes habilidades de ajuste não linear de vários modelos de previsão, de modo a descrever melhor as flutuações reais de WS/WP.

(5) Determinação de configurações ótimas para diferentes modelos de previsão. Teoricamente falando, a expressão complexa de uma DNN pode modelar qualquer relação não linear complexa. No entanto, o desempenho dos modelos de previsão baseados em DNN é bastante afetado pelas configurações do modelo, como estrutura do modelo, algoritmo de aprendizado, função de perda e assim por diante. Além disso, as características do WS/WP variam em diferentes locais. Atualmente, muitos métodos de otimização inteligentes foram adotados para determinar as configurações ótimas de modelos de previsão rasos e alcançaram melhores resultados de previsão. No entanto, é demorado e difícil encontrar configurações adequadas de modelos profundos com vários parâmetros que precisam ser ajustados.

Portanto, trabalhos futuros devem estudar regras comuns que forneçam configurações ótimas para modelos de previsão complexos com base nas características de um determinado conjunto de dados de vento. Além disso, métodos de otimização artificial mais eficientes devem ser projetados e aplicados para ajustar modelos de previsão profundos.

9. Conclusão

A exploração e utilização da energia eólica pode aliviar a crise energética e a poluição ambiental. A maneira mais direta de usar a energia eólica é transformá-la em eletricidade por meio de uma turbina eólica e integrá-la a uma rede elétrica. No entanto, a estabilidade da oferta é grandemente desafiada pela natureza estocástica do vento. Para resolver isso

problema, modelos apropriados devem ser desenvolvidos para gerar previsões precisas de WS/WP. Inspirado pela aplicação bem-sucedida de DNNs em vários domínios, este artigo revisou de forma abrangente vários modelos de DNN usados para aprimorar a precisão da previsão de WS/WP. Esta revisão identificou aplicações potenciais de modelos DNN em cada estágio da previsão de WS/WP, o que pode ajudar os pesquisadores a entender o status atual da previsão de WS/WP baseada em DNN.

Para pré-processamento de dados de vento, as abordagens mais comumente usadas são métodos de processamento de sinal e métodos de detecção de outliers, enquanto os modelos baseados em DNN raramente são aplicados especificamente para remoção de ruídos ou detecção de outliers de dados de vento. Para extração de recursos, modelos baseados em AE, modelos baseados em RBM e modelos baseados em CNN são empregados para obter recursos profundos dos dados de vento originais. Geralmente, os modelos baseados em DNN são considerados mais adequados para lidar com esse problema, pois têm duas fases de treinamento, pré-treinamento não supervisionado em camadas e ajuste fino supervisionado. A comparação de diferentes tipos de modelos baseados em CNN mostra que 1D-CNN tem a menor complexidade em oposição ao 3D-CNN que tem a maior complexidade, mas é capaz de extrair recursos espaço-temporais. Para aprendizado de relacionamento, os métodos mais populares são RNNs, incluindo LSTM, GRU, BiLSTM e BiGRU. Para tirar proveito de diferentes modelos de aprendizado de relacionamento baseados em DNN, abordagens híbridas baseadas em empilhamento e baseadas em peso foram propostas na literatura e demonstraram fazer previsões precisas de WS/WP. Eles geralmente mostram melhor desempenho de previsão do que os modelos de previsão tradicionais e baseados em DNN únicos. Na última etapa, diferentes algoritmos de otimização inteligente são empregados para obter as configurações ótimas dos modelos tradicionais de previsão, modelos baseados em DNN, bem como os pesos em modelos híbridos. No entanto, as configurações da maioria dos modelos baseados em DNN são geralmente determinadas por tentativa e erro e, devido ao alto custo computacional, apenas um pequeno número de configurações são testadas.

Para melhorar ainda mais a precisão da previsão, três desafios para a previsão WS/WP foram analisados neste artigo. Com a esperança de resolver esses desafios, cinco tendências futuras para melhorar o desempenho da previsão foram resumidas. Esta revisão fornece orientação para pesquisadores dedicados ao design de modelos ótimos de previsão de WS/WP.

Declaração de interesse concorrente

Os autores declaram que não conhecem interesses financeiros concorrentes ou relações pessoais que possam ter influenciado o trabalho relatado neste artigo.

Agradecimentos

O trabalho é apoiado pela National Natural Science Foundation of China sob Grant 62006250, o Key R&D Program of Hunan Province of China sob o Project 2020WK2007, a National Natural Science Foundation of China sob Grant 61732011, e a Natural Science Foundation of Hunan Province, China, sob Concessão 2021JJ40792.

Referências

- [1] Wang Y, Hu Q, Li L, Foley AM, Srinivasan D. Abordagens para modelagem de curvas de energia eólica: Uma revisão e discussão. *Renovar Sustain Energy Rev* 2019;116:109422.
- [2] Wang J, Song Y, Liu F, Hou R. Análise e aplicação de modelos de previsão na integração de energia eólica: Uma revisão de modelos de previsão de velocidade do vento multi-passo à frente. *Renovar Sustain Energy Rev* 2016;60:960–81.
- [3] Wang Y, Li Y, Runmin Z, Foley AM, Alkez D, Song D, Hu Q, Srinivasan D. Modelos de regressão multispline heteroscedástica esparsa para modelagem de curva de potência de turbina eólica. *IEEE Trans Sustain Energy* 2020;12(1):191–201.
- [4] Liu H, Chen C. Estratégias de processamento de dados em modelos e aplicações de previsão de energia eólica: Uma revisão abrangente. *Appl Energy* 2019;249:392–408.
- [5] Yang J, Fang L, Song D, Su M, Yang X, Huang L, Joo YH. Revisão da estratégia de controle do sistema de guinada de grandes turbinas eólicas de eixo horizontal. *Energia Eólica* 2021;24:97–115.

- [6] Lee J, Zhao F, Dutton A, Backwell B, Fiestas R, Qiao L, Balachandran N, Lim S, Liang W, Clarke E, et al. Relatório global de energia eólica 2019. Bruxelas: Conselho Global de Energia Eólica (GWEC); 2020.
- [7] IRENA. Estatísticas de capacidade renovável 2021. Abu Dhabi: Agência Internacional de Energia Renovável; 2021.
- [8] Wang Y, Hu Q, Srinivasan D, Wang Z. Modelagem de curva de energia eólica e previsão de energia eólica com dados inconsistentes. *IEEE Trans Sustain Energy* 2018;10(1):16–25.
- [9] Chen C, Duan S, Cai T, Liu B, Hu G. Sistema inteligente de gerenciamento de energia para operação econômica ideal da microrrede. *IET Renew Power Gener* 2011;5(3):258–67.
- [10] Sabzehgar R, Amirhosseini DZ, Rasouli M. Previsão de energia solar para uma microrrede inteligente residencial baseada em previsões numéricas do tempo usando métodos de inteligência artificial. *J Build Eng* 2020;32:101629.
- [11] Soman SS, Zarepour H, Malik O, Mandal P. Uma revisão da energia eólica e métodos de previsão da velocidade do vento com diferentes horizontes temporais. In: Simpósio de energia norte-americano (naps), 2010. IEEE; 2010, pág. 1–8.
- [12] Jung J, Broadwater RP. Situação atual e avanços futuros para a velocidade do vento e previsão de energia. *Renovar Sustain Energy Rev* 2014;31:762–77.
- [13] Yan J, Ouyang T. Previsão avançada de energia eólica com base em erro baseado em dados correção. *Energy Convers Manage* 2019;180:302–11.
- [14] Wei, Yi-Ming, Zhao, Weigang, Zhongyue. Previsão de velocidade do vento com um dia de antecedência : uma abordagem baseada em reamostragem. *Appl Energy* 2016;178:886–901.
- [15] Riahi G, Abedi M. Previsão de velocidade do vento de curto prazo para aplicações de turbinas eólicas usando o método de previsão linear. *Renovar Energia* 2008;33(1):35–41.
- [16] Zhang J, Yan J, Infield D, Liu Y, Lien Fs. Previsão de curto prazo e análise de incerteza de energia de turbinas eólicas com base em rede de memória de longo prazo e modelo de mistura gaussiana. *Appl Energy* 2019;241:229–44.
- [17] Prósper MA, Otero-Casal C, Fernández FC, Miguez-Macho G. Previsão de energia eólica para um parque eólico terrestre real em terreno complexo usando simulações WRF de alta resolução. *Renovar Energia* 2019;135:674–86.
- [18] Yan J, Zhang H, Liu Y, Han S, Li L, Lu Z. Previsão da alta penetração da energia eólica em múltiplas escalas usando mapeamento multi-para-multi. *IEEE Trans Power Syst* 2018;33(3):3276–84.
- [19] Yu Y, Han X, Yang M, Yang J. Previsão probabilística da energia eólica regional com base na regressão quantílica espaço-temporal. *IEEE Trans Ind Appl* 2020;56:6117–27.
- [20] Santamaría-Bonfil G, Reyes-Ballesteros A, Gershenson C. Previsão da velocidade do vento para parques eólicos: Um método baseado em regressão vetorial de suporte. *Renovar Energia* 2016;85:790–809.
- [21] Liu H, Chen C, Lv X, Wu X, Liu M. Previsão determinística de energia eólica: Uma revisão de preditores inteligentes e métodos auxiliares. *Energy Convers Manage* 2019;195:328–45.
- [22] Wang H, Wang G, Li G, Peng J, Liu Y. Rede de crença profunda baseada na abordagem de previsão da velocidade do vento determinística e probabilística. *Appl Energy* 2016;182:80–93.
- [23] Zhao X, Jiang N, Liu J, Yu D, Chang J. Velocidade média do vento de curto prazo e previsões de desvio padrão turbulento com base na rede neural convolucional unidimensional e no método de integração para a estrutura probabilística. *Energy Convers Manage* 2020;203:112239.
- [24] Yuan X, Chen C, Jiang M, Yuan Y. Intervalo de previsão de energia eólica usando o modelo LSTM baseado em distribuição Beta otimizada para parâmetros. *Appl Soft Comput* 2019;82:105550.
- [25] Hu J, Heng J, Wen J, Zhao W. Previsão determinística e probabilística da velocidade do vento com estratégia de reconstrução de ruído e algoritmo baseado em regressão quantílica. *Renovar Energia* 2020;162:1208–26.
- [26] Banik A, Behera C, Sarathkumar TV, Goswami AK. Previsão de energia eólica incerta usando intervalo de previsão baseado em LSTM. *IET Renovar Gerador de Energia* 2020;14(14):2657–67.
- [27] Zhang Y, Wang J, Wang X. Revisão sobre previsão probabilística de energia eólica geração. *Renovar Sustain Energy Rev* 2014;32:255–70.
- [28] Khosravi A, Nahavandi S. Intervalos de predição não paramétricos combinados para geração de energia eólica. *IEEE Trans Sustain Energy* 2013;4(4):849–56.
- [29] Khosravi A, Nahavandi S, Creighton D, Atiya AF. Revisão abrangente de intervalos de previsão baseados em rede neural e novos avanços. *IEEE Trans Neural Netw* 2011;22(9):1341–56.
- [30] Zhang Z, Qin H, Liu Y, Yao L, Yu X, Lu J, Jiang Z, Feng Z. Previsão da velocidade do vento com base na rede de memória fechada mínima de regressão quantílica e estimativa de densidade do kernel. *Energy Convers Manage* 2019;196:1395–409.
- [31] Khosravi A, Nahavandi S, Creighton D, Atiya AF. Método de estimativa de limite superior inferior para construção de intervalos de previsão baseados em redes neurais. *IEEE Trans Neural Netw* 2010;22(3):337–46.
- [32] Lei M, Shiyun L, Chuanwen J, Hongling L, Yan Z. Uma revisão sobre a previsão da velocidade do vento e da energia gerada. *Renovar Sustain Energy Rev* 2009;13(4):915–20.
- [33] Han Q, Meng F, Hu T, Chu F. Modelos híbridos não paramétricos para velocidade do vento previsão. *Energy Convers Manage* 2017;148:554–68.
- [34] Yunus K, Thiringer T, modelagem decomposta em frequência baseada em Chen P. Arima de série temporal da velocidade do vento. *IEEE Trans Power Syst* 2015;31(4):2546–56.
- [35] Kavasseri RG, Seetharaman K. Previsão da velocidade do vento no dia anterior usando f-ARIMA modelos. *Renovar Energia* 2009;34(5):1388–93.
- [36] Maatallah OA, Achuthan A, Janoyan K, Marzocca P. Previsão recursiva da velocidade do vento baseada no modelo auto-regressivo de hammerstein. *Appl Energy* 2015;145:191–7.
- [37] Wang J, Wang S, Yang W. Um novo sistema de combinação não linear para previsão da velocidade do vento a curto prazo. *Renovar Energia* 2019;143:1172–92.
- [38] Wang Y, Wang J, Wei X. Um modelo híbrido de previsão da velocidade do vento baseado na teoria da reconstrução do espaço de fase e no modelo de Markov: Um estudo de caso de parques eólicos no noroeste da China. *Energia* 2015;91:556–72.
- [39] Liu H, Mi X, Li Y. Modelo inteligente de aprendizado profundo de várias etapas para previsão de velocidade do vento com base na decomposição de modo variacional, análise de espectro singular, rede lstm e elm. *Energy Convers Manage* 2018;159:54–64.
- [40] Fu W, Wang K, Tan J, Zhang K. Uma estrutura composta acoplado seleção de recursos múltiplos, modelos de previsão compostos e uma nova estratégia de otimização de sincronização baseada em otimizador de enxame híbrido para previsão de velocidade de vento de curto prazo com vários passos à frente. *Energy Convers Manage* 2020;205:112461.
- [41] Severiano CA, e Silva PCdL, Cohen MW, Guimarães FG. Evolução de séries temporais fuzzy para previsão espaço-temporal em sistemas de energia renovável. *Renovar Energia* 2021;171:764–83.
- [42] Cassola F, Burlando M. Velocidade do vento e previsão de energia eólica através da filtragem kalman da saída do modelo numérico de previsão do tempo. *Appl Energy* 2012;99:154–66.
- [43] Song J, Wang J, Lu H. Um novo modelo combinado baseado em algoritmo de otimização avançado para previsão de velocidade do vento de curto prazo. *Appl Energy* 2018;215:643–58.
- [44] He X, Nie Y, Guo H, Wang J. Pesquisa em um novo sistema de combinação com base em algoritmo de otimização de aprendizado profundo e inteligência de enxame para previsão de velocidade do vento. *Acesso IEEE* 2020;8:51482–99.
- [45] Song L, Xie Q, He Y, Dang P. Modelo de previsão de combinação de energia eólica de ultracurto prazo baseado em MEEMD-SAE-Elman. In: 2020 IEEE 4ª conferência de tecnologia da informação, redes, eletrônica e controle de automação (itnec), Vol. 1. IEEE; 2020, pág. 1844–50.
- [46] Neshat M, Nezhad MM, Abbasnejad E, Mirjalili S, Tjernberg LB, Garcia DA, Alexander B, Wagner M. Um modelo evolutivo baseado em aprendizado profundo para previsão de velocidade do vento de curto prazo: um estudo de caso do parque eólico offshore lillgrund . *Energy Convers Manage* 2021;236:114002.
- [47] Niu Z, Yu Z, Tang W, Wu Q, Reformatar M. Previsão de energia eólica usando rede de unidades recorrentes fechadas baseadas em atenção. *Energia* 2020;196:117081.
- [48] Ding M, Zhou H, Xie H, Wu M, Nakanishi Y, Yokoyama R. Um modelo de correção de erro de velocidade do vento baseado em redes neurais de unidade recorrente fechada para previsão de energia eólica de curto prazo . *Neurocomputação* 2019;365:54–61.
- [49] Wan J, Liu J, Ren G, Guo Y, Yu D, Hu Q. Previsão para o dia seguinte da velocidade do vento com aprendizado profundo de recursos. *Int J Pattern Recognit Artif Intell* 2016;30(05):1650011.
- [50] Mezaache H, Bouzgou H. Auto-codificador com redes neurais para previsão da velocidade do vento. In: 2018 Conferência Internacional de Comunicações e Engenharia Elétrica (iccee). IEEE; 2018, pág. 1–5.
- [51] Shi J, Guo J, Zheng S. Avaliação de abordagens de previsão híbridas para séries temporais de velocidade do vento e geração de energia. *Renovar Sustain Energy Rev* 2012;16(5):3471–80.
- [52] Chen C, Liu H. Modelo dinâmico de previsão de velocidade do vento de conjunto baseado em aprendizado de reforço profundo híbrido. *Inform* 2021;48:101290.
- [53] Zhu Q, Chen J, Shi D, Zhu L, Bai X, Duan X, Liu Y. Aprendendo as correlações temporais e espaciais em conjunto: Uma estrutura unificada para a previsão da velocidade do vento. *IEEE Trans Sustain Energy* 2019;11(1):509–23.
- [54] Chen Y, Zhang S, Zhang W, Peng J, Cai Y. Modelo de correlação espaço-temporal multifatorial baseado em uma combinação de rede neural convolucional e rede neural de memória de curto prazo longa para previsão da velocidade do vento. *Energy Convers Manage* 2019;185:783–99.
- [55] Ghaderi A, Sanandaji BM, Ghaderi F. Previsão profunda: baseado em aprendizado profundo previsão espaço-temporal. 2017, pré-impressão arXiv arXiv:1707.08110.
- [56] Liang S, Nguyen L, Jin F. Uma rede de memória de longo prazo empilhada multivariável para previsão da velocidade do vento. In: 2018 IEEE International Conference on Big Data (Big Data). IEEE; 2018, pág. 4561–4.
- [57] Chen J, Zeng GQ, Zhou W, Du W, Lu KD. Previsão da velocidade do vento usando um conjunto de aprendizado não linear de previsão de séries temporais de aprendizado profundo e otimização extrema. *Energy Convers Manage* 2018;165:681–95.
- [58] Liu H, Chen C. Modelo multi-objetivo de previsão de velocidade do vento com conjunto de dados com autoencoder esparsos empilhado e correção de erros baseada em decomposição adaptativa. *Appl Energy* 2019;254:113686.
- [59] Qureshi AS, Khan A, Zameer A, Usman A. Previsão de energia eólica usando meta regressão baseada em rede neural profunda e aprendizado de transferência. *Appl Soft Comput* 2017;58:742–55.
- [60] Xiao L, Shao W, Jin F, Wu Z. Uma máquina de aprendizado extremo de kernel autoadaptável para previsão de velocidade do vento de curto prazo. *Appl Soft Comput* 2021;99:106917.
- [61] Li LL, Chang YB, Tseng ML, Liu JQ, Lim MK. Previsão de energia eólica usando um novo modelo em algoritmo de busca atômica aprimorado por máquinas de vetor de suporte de decomposição wavelet . *J Cleaner Prod* 2020;270:121817.
- [62] Chen Y, Dong Z, Wang Y, Su J, Han Z, Zhou D, Zhang K, Zhao Y, Bao Y. Estrutura de previsão de velocidade do vento de curto prazo com base no método eemd-ga-lstm sob histórico de vento em grande escala. *Energy Convers Manage* 2021;227:113559.

- [63] Hu YL, Chen L. Um modelo híbrido não linear de previsão de velocidade do vento usando rede LSTM, ELM histerético e algoritmo de evolução diferencial. *Energy Convers Manage* 2018;173:123–42.
- [64] Wang J, Yang W, Du P, Niu T. Um novo sistema híbrido de previsão da velocidade do vento baseado em um algoritmo de seno cosseno multi-objetivo recém-desenvolvido. *Energy Convers Manage* 2018;163:134–50.
- [65] Luo L, Li H, Wang J, Hu J. Projeto de um sistema combinado de previsão da velocidade do vento baseado na abordagem de otimização multiobjetivo e conjunto de decomposição. *Appl Math Model* 2021;89:49–72.
- [66] Wu C, Wang J, Chen X, Du P, Yang W. Um novo sistema híbrido baseado em otimização multiobjetivo para previsão da velocidade do vento. *Renovar Energia* 2020;146:149–65.
- [67] Landberg L. Prediktor-um sistema de previsão on-line. In: *Energia eólica para o século 21: conferência e exposição de tema especial euwec*. 2000.
- [68] Nielsen T, Madsen H. WPPT - uma ferramenta para previsão de energia eólica. In: *Prediction.Proceedings of wind power for the 21st century*, Kassel, Alemanha, 25-27 setembro 2000. 2000.
- [69] ENFOR. *Software de previsão de energia eólica: WindForTM*. 2021, <https://enfor.dk/servicos/windfor/> (Acessado em 30 de julho de 2021).
- [70] Giebel G, Landberg L, Kariniotakis G, Brownsword R. Métodos e ferramentas de software de última geração para previsão de curto prazo da produção de energia eólica. In: *European Wind Energy Conference and Exhibition* 2003. 2003.
- [71] Monteiro C, Bessa R, Miranda V, Botterud A, Wang J, Conzelmann G, et al. *Previsão de energia eólica: estado da arte* 2009. Tech. rep., Argonne, IL (Estados Unidos): Argonne National Lab.(ANL); 2009.
- [72] AleaSoft. *AleaSoft: Previsão de energia*. 2021, <https://aleasoft.com/> (Acessado em 30 de julho de 2021).
- [73] Elkin C, Witherspoon S. O aprendizado de máquina pode aumentar o valor da energia eólica. 2021, <https://deepmind.com/blog/article/machine-learning-can-boost-value-wind-energy> (Acessado em 31 de julho de 2021).
- [74] Marugán AP, Márquez FPG, Perez JMP, Ruiz-Hernández D. Um levantamento da rede neural artificial em sistemas de energia eólica. *Appl Energy* 2018;228:1822–36.
- [75] Zendejboudi A, Baseer M, Saidur R. Aplicação de modelos de máquinas vetoriais de suporte para previsão de recursos de energia solar e eólica: uma revisão. *J Cleaner Prod* 2018;199:272–85.
- [76] Tascikaraoglu A, Uzunoglu M. Uma revisão de abordagens combinadas para previsão de velocidade e potência do vento a curto prazo. *Renovar Sustain Energy Rev* 2014;34:243–54.
- [77] Xiao L, Wang J, Dong Y, Wu J. Modelos de previsão combinados para previsão de energia eólica : Um estudo de caso na China. *Renovar Sustain Energy Rev* 2015;44:271–88.
- [78] Ren Y, Suganthan P, Srikanth N. Ensemble métodos para previsão de energia eólica e solar – uma revisão de última geração. *Renovar Sustain Energy Rev* 2015;50:82–91.
- [79] Liu H, Li Y, Duan Z, Chen C. Uma revisão sobre a estrutura de otimização multiobjetivo em técnicas e aplicações de previsão de energia eólica. *Energy Convers Manage* 2020;224:113324.
- [80] Qian Z, Pei Y, Zareipour H, Chen N. Uma revisão e discussão de modelos híbridos baseados em decomposição para aplicações de previsão de energia eólica. *Appl Energy* 2019;235:939–53.
- [81] Ahmed A, Khalid M. Uma revisão sobre as aplicações selecionadas de modelos de previsão em sistemas de energia renovável. *Renovar Sustain Energy Rev* 2019;100:9–21.
- [82] Okumus I, Dinler A. Situação atual da previsão de energia eólica e um método híbrido para previsões horárias. *Energy Convers Manage* 2016;123:362–71.
- [83] Ahmad T, Zhang H, Yan B. Uma revisão sobre modelos de previsão de necessidades de energia renovável e eletricidade para redes inteligentes e edifícios. *Cidades Sustentáveis Soc* 2020;55:102052.
- [84] Khalid R, Javaid N. Uma pesquisa sobre algoritmos de otimização de hiperparâmetros de modelos de previsão em smart grid. *Cidades Sustentáveis Soc* 2020;61:102275.
- [85] Ssekulima EB, Anwar MB, Al Hinaí A, El Moursi MS. Técnicas de previsão de velocidade do vento e irradiância solar para integração aprimorada de energia renovável com a rede: uma revisão. *IET Renew Power Gener* 2016;10(7):885–989.
- [86] Afrasiabi M, Mohammadi M, Rastegar M, Afrasiabi S. Abordagem avançada de aprendizado profundo para previsão probabilística da velocidade do vento. *IEEE Trans Ind Inf* 2020;17(1):720–7.
- [87] Zhang Y, Li Y, Zhang G. Abordagem de previsão de energia eólica de curto prazo baseada em Modelo Seq2Seq usando dados NWP. *Energia* 2020;118371.
- [88] Higashiyama K, Fujimoto Y, Hayashi Y. Extração de recursos de dados NWP para previsão de energia eólica usando redes neurais convolucionais 3D. *Energia Procedia* 2018;155:350–8.
- [89] Ju Y, Sun G, Chen Q, Zhang M, Zhu H, Rehman MU. Um modelo que combina rede neural convolucional e algoritmo lightgbm para previsão de energia eólica de ultracurto prazo . Acesso IEEE 2019;1.
- [90] Harbola S, Coors V. Rede neural convolucional unidimensional arquiteturas para previsão do vento. *Energy Convers Manage* 2019;195:70–5.
- [91] Vincent P, Larochelle H, Lajoie I, Bengio Y, Manzagol PA, Bottou L. Autoencoders de remoção de ruído empilhados : Aprendendo representações úteis em uma rede profunda com um critério de remoção de ruído local. *J Mach Learn Res* 2010;11(12).
- [92] Li C, Tang G, Xue X, Saeed A, Hu X. Previsão de intervalo de velocidade do vento de curto prazo com base no modelo GRU de conjunto. *IEEE Trans Sustain Energy* 2019;11(3):1370–80.
- [93] Zu XR, Song RX. Método de previsão de energia eólica de curto prazo com base na decomposição de pacotes wavelet e GRU aprimorado. *J Phys Conf Ser* 2018;1087.
- [94] Peng Z, Peng S, Fu L, Lu B, Tang J, Wang K, Li W. Um novo modelo de conjunto de aprendizado profundo com redução de ruído de dados para previsão de velocidade do vento de curto prazo. *Energy Convers Manage* 2020;207:112524.
- [95] Jahangir H, Tayarani H, Gougheri SS, Golkar MA, Ahmadian A, Elkamel A. Abordagem de previsão baseada em aprendizado profundo em redes inteligentes com microagrupamento e rede LSTM bidirecional. *IEEE Trans Ind Electron* 2020;68(9):8298–309.
- [96] Li G, Wang Tf, Hu Fx, Liu Tc. A aplicação do algoritmo de rede neural profunda considerando extração de características meteorológicas multidimensionais na previsão da velocidade do vento a curto prazo. In: *2019 IEEE 3ª conferência de tecnologia da informação, redes, eletrônica e controle de automação (itnec)*. IEEE; 2019, pág. 2432–7.
- [97] Yang M, Shi C, Liu H. Previsão de energia eólica para o dia seguinte com base no agrupamento de curvas de potência equivalentes. *Energia* 2021;218:119515.
- [98] Liu X, Zhang H, Kong X, Lee KY. Previsão da velocidade do vento usando neural profundo rede com seleção de recursos. *Neurocomputação* 2020;397:393–403.
- [99] Wang J, Xiong S. Um modelo de previsão híbrido baseado na detecção de valores discrepantes e séries temporais difusas – um estudo de caso no parque eólico de hainan da China. *Energia* 2014;76:526–41.
- [100] Wang Y, Wang H, Srinivasan D, Hu Q. Regressão funcional robusta para previsão da velocidade do vento com base na aprendizagem bayesiana esparsa. *Renovar Energia* 2019;132:43–60.
- [101] Li Y, Wang Y, Chen Z, Zou R. Bayeslan máquina de aprendizagem multi-extrema robusta. Sistema baseado em conhecimento 2020;210:106468.
- [102] Guo Z, Zhao W, Lu H, Wang J. Previsão de várias etapas para a velocidade do vento usando um modelo de rede neural artificial baseado em EMD modificado. *Renovar Energia* 2012;37(1):241–9.
- [103] Liu Z, Jiang P, Zhang L, Niu X. Um modelo de previsão combinado para séries temporais: Aplicação à previsão da velocidade do vento a curto prazo. *Appl Energy* 2020;259:114137.
- [104] Osório G, Matias J, Catalão J. Previsão de energia eólica de curto prazo usando sistema de inferência neuro-fuzzy adaptativo combinado com otimização evolutiva de enxame de partículas, transformada wavelet e informação mútua. *Renovar Energia* 2015;75:301–7.
- [105] Liu H, Tian Hq, Li Yf. Comparação de novos híbridos FEEMD-MLP, FEEMD-ANFIS, wavelet packet-MLP e wavelet packet-ANFIS para previsões de velocidade do vento. *Energy Convers Manage* 2015;89:1–11.
- [106] Liu H, Mi X, Li Y. Comparação de duas novas abordagens inteligentes de previsão da velocidade do vento com base na decomposição de pacotes wavelet, decomposição em modo empírico de conjunto completo com ruído adaptativo e redes neurais artificiais. *Energy Convers Manage* 2018;155:188–200.
- [107] Zhang D, Peng X, Pan K, Liu Y. Uma nova previsão de velocidade do vento baseada em decomposição híbrida e máquina de aprendizado extremo robusto de outlier sequencial online. *Energy Convers Manage* 2019;180:338–57.
- [108] Wu Z, Xiao L. Uma estrutura híbrida baseada em decomposição secundária com análise meteorológica para previsão de velocidade do vento determinística e probabilística. *Appl Soft Comput* 2019;85:105799.
- [109] Liu H, Mi Xw, Li Yf. Método de previsão da velocidade do vento baseado em estratégia de aprendizado profundo usando transformada wavelet empírica, rede neural de memória de longo prazo e rede neural elman. *Energy Convers Manage* 2018;156:498–514.
- [110] Han L, Jing H, Zhang R, Gao Z. Previsão de energia eólica baseada em rede de memória de longo prazo melhorada. *Energia* 2019;189:116300.
- [111] Moreno SR, da Silva RG, Mariani VC, dos Santos Coelho L. Previsão da velocidade do vento multi-etapas com base no modelo híbrido de decomposição multi-estágio e rede neural de memória de longo prazo. *Energy Convers Manage* 2020;213:112869.
- [112] Sun N, Zhou J, Chen L, Jia B, Tayyab M, Peng T. Um modelo dinâmico adaptativo de previsão de velocidade do vento de curto prazo usando decomposição secundária e uma máquina de aprendizagem extrema regularizada melhorada. *Energia* 2018;165:939–57.
- [113] Peng T, Zhang C, Zhou J, Nazir MS. Modelo de conjunto RELM baseado em aprendizado de correlação negativa integrado ao OVMD para previsão de velocidade do vento em várias etapas. *Renovar Energia* 2020;156:804–19.
- [114] Mi X, Liu H, Li Y. Modelo de previsão da velocidade do vento usando análise de espectro singular, decomposição de modo empírico e máquina de vetor de suporte convolucional. *Energy Convers Manage* 2019;180:196–205.
- [115] Santhosh M, Venkaiah C, Kumar DV. Método de rede neural wavelet adaptável baseado em decomposição de modo empírico ensemble para previsão da velocidade do vento. *Energy Convers Manage* 2018;168:482–93.
- [116] Qu Z, Mao W, Zhang K, Zhang W, Li Z. Previsão da velocidade do vento em várias etapas com base em uma técnica de decomposição híbrida e uma rede neural de retropropagação aprimorada. *Renovar Energia* 2019;133:919–29.
- [117] Zhang W, Qu Z, Zhang K, Mao W, Ma Y, Fan X. Um modelo combinado baseado em CEEMDAN e algoritmo de polinização de flores modificado para previsão da velocidade do vento. *Energy Convers Manage* 2017;136:439–51.
- [118] Li C, Zhu Z, Yang H, Li R. Um sistema híbrido inovador para previsão de velocidade do vento baseado em esquema de pré-processamento fuzzy e otimização multi-objetivo. *Energia* 2019;174:1219–37.
- [119] Dong Q, Sun Y, Li P. Um novo modelo de previsão baseado em uma estratégia de processamento híbrido e uma rede neural difusa linear local otimizada para fazer previsão de energia eólica: Um estudo de caso de parques eólicos na China. *Renovar Energia* 2017;102:241–57.

- [120] Moreno SR, dos Santos Coelho L. Abordagem de previsão da velocidade do vento baseada em análise de espectro singular e sistema de inferência neurofuzzy adaptativo. *Renovar Energia* 2018;126:736–54.
- [121] Niu T, Wang J, Zhang K, Du P. Previsão da velocidade do vento com vários passos à frente com base na seleção de recursos ideais e um algoritmo de morcego modificado com a estratégia de cognição. *Renovar Energia* 2018;118:213–29.
- [122] Yan X, Liu Y, Xu Y, Jia M. Previsão de várias etapas para a velocidade do vento diurna com base no modelo híbrido de aprendizado profundo com decomposição de espectro singular aprimorada. *Energy Convers Manage* 2020;225:113456.
- [123] Liu H, Tian Hq, Liang Xf, Li Yf. Abordagem de previsão da velocidade do vento usando algoritmo de decomposição secundária e redes neurais de Elman. *Appl Energy* 2015;157:183–94.
- [124] Xiang L, Li J, Hu A, Zhang Y. Previsão determinística e probabilística de várias etapas para a velocidade do vento de curto prazo com base na decomposição secundária e um método de aprendizado profundo. *Energy Convers Manage* 2020;220:113098.
- [125] Mi XW, Liu H, Li YF. Método de previsão de velocidade do vento usando wavelet, máquina de aprendizado extremo e algoritmo de correção de outlier. *Energy Convers Manage* 2017;151:709–22.
- [126] Wu Z, Xia X, Xiao L, Liu Y. Modelo combinado com seleção de modelo de decomposição secundária e seleção de amostra para previsão de energia eólica em várias etapas. *Appl Energy* 2020;261:114345.
- [127] Moreno SR, Mariani VC, Coelho LD. Decomposição híbrida multiestágio com modelo paramétrico aplicado à previsão da velocidade do vento no nordeste brasileiro. *Renovar Energia* 2021;164:1508–26.
- [128] Peng T, Zhou J, Zhang C, Zheng Y. Previsão de velocidade do vento multi-passo à frente usando um modelo híbrido baseado na técnica de decomposição de dois estágios e máquina de aprendizado extremo adaboost. *Energy Convers Manage* 2017;153(dec.):589–602.
- [129] Liu H, Duan Z, Wu H, Li Y, Dong S. Modelos de previsão de velocidade do vento baseados na decomposição de dados, seleção de recursos e método de grupo de rede de manipulação de dados. *Medição* 2019;148:106971.
- [130] Wang J, Qin S, Zhou Q, Jiang H. Previsão de velocidade de vento de médio prazo utilizando modelos híbridos para três locais diferentes em xinjiang, China. *Renovar Energia* 2015;76:91–101.
- [131] Ouyang T, Huang H, He Y, Tang Z. Previsão de séries temporais de energia eólica caótica através de modos de comutação de dados. *Renovar Energia* 2020;145:270–81.
- [132] Wang S, Li B, Li G, Yao B, Wu J. Previsão de energia eólica de curto prazo com base na limpeza de dados multidimensionais e reconfiguração de recursos. *Appl Energy* 2021;292:116851.
- [133] He Y, Li H, Wang S, Yao X. Análise de incerteza de previsão de densidade de probabilidade de energia eólica com base em interpolação spline cúbica e regressão quantílica vetorial de suporte. *Neurocomputação* 2021;430:121–37.
- [134] Vincent P, Larochelle H, Bengio Y, Manzagol PA. Extrair e compor recursos robustos com autoencoders de redução de ruído. In: *Proceedings of the 25th international conference on machine learning*, 2008, p. 1096–103.
- [135] Xiaosheng P, Bo W, Fan Y, Gaofeng F, Zheng W, Kai C. Uma abordagem de aprendizado profundo para previsão de energia eólica baseada em codificadores automáticos de denoising empilhados otimizados por algoritmo de morcego. In: 2018 China International Conference on Electricity Distribution (CICED). IEEE; 2018, pág. 945–8.
- [136] Jahangir H, Golkar MA, Alhameli F, Mazouz A, Ahmadian A, Elkamel A. Estrutura de previsão de velocidade do vento de curto prazo com base em codificadores automáticos de remoção de ruído empilhados com ANN aproximado. *Sustain Energy Technol Assessment* 2020;38:100604.
- [137] Wu YX, Wu QB, Zhu JQ. Previsão de velocidade do vento orientada por dados usando extração profunda de recursos e LSTM. *IET Renew Gener* 2019;13(12):2062–9.
- [138] Feng C, Cui M, Hodge BM, Zhang J. Uma metodologia multimodelo orientada por dados com seleção profunda de recursos para previsão de vento de curto prazo. *Appl Energy* 2017;190:1245–57.
- [139] Song Z, Tang Y, Ji J, Todo Y. Avaliando um modelo de neurônio dendritico para previsão da velocidade do vento. *Sistema baseado em conhecimento* 2020;201:106052.
- [140] Hu R, Hu W, Gökmen N, Li P, Huang Q, Chen Z. Previsão de velocidade do vento de alta resolução com base na reconstrução do espaço de fase decomposta wavelet e mapa auto-organizado. *Renovar Energia* 2019;140:17–31.
- [141] Liu D, Niu D, Wang H, Fan L. Previsão de velocidade do vento de curto prazo usando transformada wavelet e máquinas de vetor de suporte otimizadas por algoritmo genético. *Renovar Energia* 2014;62:592–7.
- [142] Zhang C, Wei H, Zhao J, Liu T, Zhu T, Zhang K. Previsão de velocidade do vento de curto prazo usando decomposição de modo empírico e seleção de recursos. *Renovar Energia* 2016;96:727–37.
- [143] Memarzadeh G, Keynia F. Um novo método de previsão de velocidade do vento de curto prazo baseado na rede neural LSTM ajustada e conjuntos de entrada ótimos. *Energy Convers Manage* 2020;213:112824.
- [144] Amjady N, Keynia F, Zareipour H. Previsão de energia eólica por um novo mecanismo de previsão composto de rede neural híbrida modificada e otimização aprimorada de enxame de partículas. *IEEE Trans Sustain Energy* 2011;2(3):265–76.
- [145] Li S, Wang P, Goel L. Previsão de energia eólica usando conjuntos de redes neurais com seleção de recursos. *IEEE Trans Sustain Energy* 2015;6(4):1447–56.
- [146] Huang N, Xing E, Cai G, Yu Z, Qi B, Lin L. Previsão de velocidade do vento de curto prazo com base na seleção de recursos de baixa redundância. *Energias* 2018;11(7):1638.
- [147] Dalto M, Matusko J, Vašák M. Redes neurais profundas para previsão de vento de curto prazo. In: 2015 IEEE conferência internacional de tecnologia industrial (iciti). IEEE; 2015, pág. 1657–63.
- [148] Zhu S, Yuan X, Xu Z, Luo X, Zhang H. Modelo de mistura gaussiana acoplado a redes neurais recorrentes para previsão de intervalo de velocidade do vento. *Energy Convers Manage* 2019;198:111772.
- [149] Kong X, Liu X, Shi R, Lee KY. Previsão da velocidade do vento usando máquinas de vetor de suporte reduzido com seleção de recursos. *Neurocomputing* 2015;169:449–56.
- [150] Liu Z, Hajiali M, Torabi A, Ahmadi B, Simoes R. Novo modelo de previsão baseado em transformada wavelet melhorada, seleção de recursos informativos e máquina de vetor de suporte híbrido na previsão de energia eólica. *J Ambiente Intell Humaniz Comput* 2018;9(6):1919–31.
- [151] Sun S, Qiao H, Wei Y, Wang S. Uma nova abordagem dinâmica integrada para a previsão da velocidade do vento. *Appl Energy* 2017;197(JUL.1):151–62.
- [152] Mezaache H, Bouzgou H, Raymond C. Análise de componentes principais do kernel com máquinas de aprendizado extremo para previsão da velocidade do vento. In: Sétimo Congresso Internacional de Energias Renováveis, IREC2016. 2016.
- [153] Leng H, Li X, Zhu J, Tang H, Zhang Z, Ghadimi N. Um novo método de previsão de energia eólica baseado em transformadas de ridgelet, seleção de características híbridas e previsão de malha fechada. *Informe de Eng Adv* 2018;36:20–30.
- [154] Wang L, Zhang Z, Chen J. Previsão de preços de eletricidade de curto prazo com autoencoders de denoising empilhados. *IEEE Trans Power Syst* 2017;32(4):2673–81.
- [155] Khodayar M, Teshnehlab M. Rede neural profunda robusta para previsão da velocidade do vento. In: 2015 4º Congresso Conjunto Iraniano de Sistemas Fuzzy e Inteligentes (cfis). IEEE; 2015, pág. 1–5.
- [156] Jaseena K, Kooroor BC. Um modelo híbrido de previsão de velocidade do vento usando autoencoder empilhado e LSTM. *J Renew Sustain Energy* 2020;12(2):023302.
- [157] Khodayar M, Kaynak O, Khodayar ME. Arquitetura neural profunda áspere para previsão da velocidade do vento a curto prazo. *IEEE Trans Ind Inf* 2017;13(6):2770–9.
- [158] Jiao R, Huang X, Ma X, Han L, Tian W. Um modelo que combina codificador automático empilhado e algoritmo de retropropagação para previsão de energia eólica de curto prazo. *Acesso IEEE* 2018;6:17851–8.
- [159] Liu JN, Hu Y, You JJ, Chan PW. Representação de recursos baseada em rede neural profunda para previsão do tempo. In: Anais da Conferência Internacional de Inteligência Artificial (ICAI). O Comitê Diretor do Congresso Mundial de Ciência da Computação, Computação; 2014, pág. 1.
- [160] Yin Q, Duan B, Shen M, Qu X. Método de localização e detecção de falhas baseado em autoencoder esparsos empilhado para conversores modulares de cinco níveis. In: IECN 2017-43rd Annual Conference of the IEEE Industrial Electronics Society. IEEE; 2017, pág. 1580–5.
- [161] Saufi SR, Ahmad ZAB, Leong MS, Lim MH. Diagnóstico de falhas da caixa de engrenagens usando um modelo de aprendizado profundo com amostra de dados limitada. *IEEE Trans Ind Inf* 2020;16(10):6263–71.
- [162] Liu M, Cao Z, Zhang J, Wang L, Huang C, Luo X. Previsão da velocidade do vento a curto prazo com base no modelo jaya-SVM. *Int J Electr Power Energy Syst* 2020;121:106056.
- [163] Hinton GE. Um guia prático para treinar máquinas Boltzmann restritas. Dentro: *Redes neurais: truques do comércio*. Springer; 2012, pág. 599–619.
- [164] Zhang CY, Chen CP, Gan M, Chen L. Máquina de Boltzmann profunda preditiva para previsão de velocidade do vento em vários períodos. *IEEE Trans Sustain Energy* 2015;6(4):1416–25.
- [165] Lin KP, Pai PF, Ting YJ. Redes de crenças profundas com algoritmos genéticos na previsão da velocidade do vento. *Acesso IEEE* 2019;7:99244–53.
- [166] Hu Y, Liu JN, You J, Chan PW. Rede neural profunda contínua baseada em RBM para previsão de velocidade do vento em Hong Kong. In: Anais da Conferência Internacional de Processamento de Imagens, Visão Computacional e Reconhecimento de Padrões (IPCV). O Comitê Diretor do Congresso Mundial de Ciência da Computação, Computação; 2015, pág. 368.
- [167] Tao Y, Chen H, Qiu C. Previsão de energia eólica e recurso de padrão baseado no método de aprendizado profundo. In: 2014 IEEE Asia-Pacific Power and Energy Engineering Conference (APPEEC). IEEE; 2014, pág. 1–4.
- [168] Wang K, Qi X, Liu H, Song J. Abordagem de cluster k-means baseada em rede de crença profunda para previsão de energia eólica de curto prazo. *Energia* 2018;165:840–52.
- [169] Jiajun H, Chuanjin Y, Yongle L, Huoyue X. Previsão de vento de prazo ultracurto com transformada wavelet, rede de crença profunda e aprendizado de conjunto. *Energy Convers Manage* 2020;205:112418.
- [170] Zhang Y, Le J, Liao X, Zheng F, Li Y. Um novo modelo de previsão de combinação para energia eólica integrando máquina de vetor de suporte mínimo quadrado, rede de crença profunda, análise de espectro singular e hashing sensível à localidade. *Energia* 2019;168:558–72.
- [171] Zhu A, Li X, Mo Z, Wu R. Previsão de energia eólica baseada em uma rede neural convolucional. In: 2017 conferência internacional sobre circuitos, dispositivos e sistemas (iccds). IEEE; 2017, pág. 131–5.
- [172] He Z, Shao H, Zhong X, Zhao X. CNNs de transferência de conjunto acionadas por sinais multicanal para diagnóstico de falhas de máquinas rotativas em condições de trabalho cruzadas. *Sistema baseado em conhecimento* 2020;207:106396.
- [173] Hong YY, Rioflorida CLPP. Uma rede neural híbrida baseada em aprendizado profundo para Previsão de energia eólica com 24 horas de antecedência. *Appl Energy* 2019;250:530–9.
- [174] Liu H, Mi X, Li Y. Modelo de previsão de velocidade do vento baseado em aprendizado profundo inteligente usando decomposição de pacotes wavelet, rede neural convolucional e rede de memória de longo prazo convolucional. *Energy Convers Manage* 2018;166:120–31.
- [175] Zhang G, Liu D. Rede convolucional de unidades recorrentes com múltiplos métodos de decomposição para previsão de velocidade do vento a curto prazo. *Energy Convers Manage* 2020;226:113500.

- [176] Hong YY, Satriani TRA. Previsão de velocidade do vento espaço-temporal com antecedência usando rede neural de aprendizado profundo baseada em design robusto. *Energia* 2020;118441.
- [177] Yu R, Liu Z, Li X, Lu W, Yu M, Wang J, Li B. Aprendizagem de cena: Redes convolucionais profundas para previsão de energia eólica incorporando turbinas no espaço da grade. *Energia Aplicada* 2019;238:249–57.
- [178] Wang Hz, Li Gq, Wang Gb, Peng Jc, Jiang H, Liu Yt. Abordagem de conjunto baseada em aprendizado profundo para previsão probabilística de energia eólica. *Appl Energy* 2017;188:56–70.
- [179] Liu Y, Qin H, Zhang Z, Pei S, Jiang Z, Feng Z, Zhou J. Previsão probabilística espaço-temporal da velocidade do vento com base em um modelo de aprendizado profundo Bayesian variacional. *Appl Energy* 2020;260:114259.
- [180] Mehrkanon S. Aprendizagem de representação compartilhada profunda para previsão de elementos climáticos . Sistema Baseado em Conhecimento 2019;179(SEP.1):120–8.
- [181] Trebing K, Mehrkanon S. Previsão da velocidade do vento usando redes neurais convolucionais multidimensionais. In: 2020 IEEE Symposium Series on Computational Intelligence (ssci). IEEE; 2020, pág. 713–20.
- [182] Notchenko A, Kapushev E, Burnaev E. Redes neurais convolucionais 3D esparsas para recuperação de forma em larga escala. In *Proceedings of neural information processing systems*, Vol. 29, 2016.
- [183] Kiranyaz S, Avci O, Abdeljaber O, Ince T, Gabbouj M, Inman DJ. Redes neurais convolucionais 1d e aplicações: Uma pesquisa. *Processo de sinal do sistema mecânico* 2021;151:107398.
- [184] Hochreiter S, Schmidhuber J. Longa memória de curto prazo. *Computação Neural* 1997;9(8):1735–80.
- [185] Araya IA, Valle C, Allende H. Um modelo multi-escala baseado na memória de longo prazo para a previsão da velocidade do vento por hora. *Reconhecimento de padrões Lett* 2019.
- [186] Yu R, Gao J, Yu M, Lu W, Xu T, Zhao M, Zhang J, Zhang R, Zhang Z. Lstm-efg para previsão de energia eólica com base em recursos de correlação sequencial. *Sistema de Computação do Futuro Gerador* 2019;93:33–42.
- [187] Li Y, Wu H, Liu H. Previsão da velocidade do vento em várias etapas usando decomposição EWT , computação principal LSTM, computação subordinada RELM e reconstrução IEWT. *Energy Convers Manage* 2018;167:203–19.
- [188] Han L, Zhang R, Wang X, Bao A, Jing H. Previsão de energia eólica em várias etapas com base no VMD-LSTM. *IET Renovar Gerador de Energia* 2019;13(10):1690–700.
- [189] Juijie W, Yaning L. Previsão de velocidade do vento de vários passos à frente com base na extração de recursos ideal, rede neural de memória de longo prazo e estratégia de correção de erros. *Appl Energy* 2018;230:429–43.
- [190] Chen MR, Zeng GQ, Lu KD, Weng J. Um método de combinação não linear de duas camadas para previsão de velocidade do vento de curto prazo com base em ELM, ENN e LSTM. *IEEE Internet Things J* 2019;6(4):6997–7010.
- [191] Schuster M, Paliwal K. Redes neurais recorrentes bidirecionais. *IEEE Trans Signal Process* 2002;45(11):2673–81.
- [192] Biswas S, Sinha M. Performances de modelos de aprendizado profundo para vento do oceano Índico previsão de velocidade. *Modelo Earth Syst Environ* 2020.
- [193] Saeed A, Li C, M dinamarquês, Rubaiee S, Tang G, Gan Z, Ahmed A. Modelo LSTM bidirecional híbrido para previsão de intervalo de velocidade do vento de curto prazo. *Acesso IEEE* 2020;8:182283–94.
- [194] Xiang J, Qiu Z, Hao Q, Cao H. Previsão da velocidade do vento em escala múltipla com base em WT-bi-LSTM. *MATEC Web Conf* 2020;309(1):05011.
- [195] Chen G, Li L, Zhang Z, Li S. Previsão de velocidade do vento de curto prazo com principal preditor subordinado baseado em conv-LSTM e BPNN melhorado. *Acesso IEEE* 2020;8:67955–73.
- [196] Shi X, Chen Z, Wang H, Yeung DY, Wong WK, Woo Wc. Rede LSTM convolucional : Uma abordagem de aprendizado de máquina para previsão de precipitação. *Sistema de Processo de Inf Neural Adv* 2015;28:802–10.
- [197] Wilms H, Cupelli M, Monti A, Gross T. Explorando dependências espaço-temporais para previsões de energia eólica baseadas em RNN. In: IEEE pes gtd Grand International Conference and Exposition Asia. IEEE; 2019, pág. 921–6.
- [198] Ibrahim M, Alsheikh A, Al-Hindawi Q, Al-Dahidi S, Elmoaqet H. Previsão da velocidade do vento de curto prazo usando algoritmos baseados em aprendizado artificial. *Computar Intel Neuroence* 2020;2020(12).
- [199] Sun Z, Zhao M. Previsão de energia eólica de curto prazo com base na decomposição VMD , redes de conversão e análise de erros. *Acesso IEEE* 2020;8:134422–34.
- [200] Woo S, Park J, Park J. Previsão de potência de turbina eólica e saídas de carga por modelo LSTM convolucional multitarefa. In: 2018 IEEE power & energy society assembly geral (pesgm). IEEE; 2018, pág. 1–5.
- [201] Qin Y, Li K, Liang Z, Lee B, Zhang F, Gu Y, Zhang L, Wu F, Rodriguez D. Modelo de previsão híbrido baseado em rede de memória de longo prazo e rede neural de aprendizado profundo para sinal de vento. *Appl Energy* 2019;236:262–72.
- [202] Cho K, Van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y. Aprendendo representações de frases usando o codificador-decodificador RNN para tradução automática estatística. In: EMNLP 2014 - 2014 conferência sobre métodos empíricos em processamento de linguagem natural, anais da conferência. 2014, pág. 1724–34.
- [203] Kisvari A, Lin Z, Liu X. Previsão de energia eólica – um método orientado a dados juntamente com rede neural recorrente fechada. *Renovar Energia* 2021;163:1895–909.
- [204] Liu H, Mi X, Li Y, Duan Z, Xu Y. Aprendizado profundo de velocidade do vento inteligente baseado em modelo de previsão de várias etapas usando análise de espectro singular, rede de unidade recorrente fechada convolucional e regressão vetorial de suporte. *Renovar Energia* 2019;143:842–54.
- [205] Yu C, Li Y, Bao Y, Tang H, Zhai G. Uma nova estrutura para previsão de velocidade do vento com base em redes neurais recorrentes e máquina de vetor de suporte. *Energy Convers Manage* 2018;178:137–45.
- [206] Wang R, Li C, Fu W, Tang G. Método de aprendizado profundo baseado em unidade recorrente fechada e decomposição de modo variacional para previsão de intervalo de energia eólica de curto prazo . *IEEE Trans Neural Netw Learn Syst* 2020;PP(99):1–14.
- [207] Zhang H, Liu Y, Yan J, Han S, Li L, Long Q. Rede de densidade de mistura profunda aprimorada para previsão probabilística de energia eólica regional. *IEEE Trans Power Syst* 2020;35(4):2549–60.
- [208] Khodayar M, Wang J. Rede neural profunda do gráfico espaço-temporal para a previsão da velocidade do vento a curto prazo. *IEEE Trans Sustain Energy* 2018;10(2):670–81.
- [209] Wen H, Gu J, Ma J, Jin Z. Previsão probabilística de energia eólica através de intervalos de previsão baseados em aprendizado profundo Bayesiano. In: 2019 IEEE 17th international conference on industrial informática (indin), Vol. 1. IEEE; 2019, pág. 1091–6.
- [210] Hu H, Wang L, Lv SX. Previsão de consumo de energia e geração de energia eólica usando rede de estado de eco profundo. *Renovar Energia* 2020;154:598–613.
- [211] Liu H, Yu C, Wu H, Duan Z, Yan G. Um novo modelo de aprendizado de reforço profundo de conjunto híbrido para previsão de curto prazo da velocidade do vento. *Energia* 2020;202:117794.
- [212] Yang HF, Chen Y-PP. Aprendizado de representação com máquinas de aprendizado extremo e decomposição de modo empírico para métodos de previsão de velocidade do vento. *Inteligência Artificial* 2019;277:103176.
- [213] Zhang S, Chen Y, Xiao J, Zhang W, Feng R. Modelo híbrido de previsão de velocidade do vento baseado em abordagem de decomposição secundária de dados multivariados e algoritmo de aprendizado profundo com mecanismo de atenção. *Renovar Energia* 2021;174:688–704.
- [214] Woo S, Park J, Park J, Manuel L. Previsão de resposta de turbina de curto prazo baseada em campo eólico por Istms convolucionais dilatados empilhados. *IEEE Trans Sustain Energy* 2019;11(4):2294–304.
- [215] Yan H, Wu Z. Um modelo híbrido de previsão de energia eólica de curto prazo combinando processamento de dados, otimização de vários parâmetros e estratégia de distribuição de modelos multi-inteligentes . *Acesso IEEE* 2020;8:227126–40.
- [216] Pei S, Qin H, Zhang Z, Yao L, Wang Y, Wang C, Liu Y, Jiang Z, Zhou J, Yi T. Método de previsão da velocidade do vento baseado na transformação wavelet empírica e nova atualização da célula a longo prazo rede de memória. *Energy Convers Manage* 2019;196:779–92.
- [217] Liang T, Xie G, Fan S, Meng Z. Um modelo combinado baseado em ceemdan, entropia de permutação, rede de unidades recorrentes fechadas e um algoritmo de morcego aprimorado para previsão de velocidade do vento. *Acesso IEEE* 2020;8:165612–30.
- [218] Liu MD, Ding L, Bai YL. Aplicação de modelo híbrido baseado em decomposição de modo empírico, novas redes neurais recorrentes e arima para previsão de velocidade do vento. *Energy Convers Manage* 2021;233:113917.
- [219] Jiang P, Liu Z, Niu X, Zhang L. Um sistema de previsão combinado baseado em método estatístico, redes neurais artificiais e métodos de aprendizado profundo para previsão de velocidade do vento de curto prazo. *Energia* 2021;217:119361.
- [220] Chen Y, He Z, Shang Z, Li C, Li L, Xu M. Um novo modelo combinado baseado na rede de estado de eco para previsão de velocidade do vento multi-passo à frente: Um estudo de caso de nrel. *Energy Convers Manage* 2019;179:13–29.
- [221] He Z, Chen Y, Shang Z, Li C, Li L, Xu M. Um novo modelo de previsão da velocidade do vento baseado em janela móvel e algoritmo de otimização de enxame de partículas multi-objetivo . *Appl Math Model* 2019;76:717–40.
- [222] Wei D, Wang J, Niu X, Li Z. Sistema de previsão de velocidade do vento baseado em unidades recorrentes fechadas e redes neurais de pico convolucional. *Appl Energy* 2021;292:116842.
- [223] Al-Shabandar R, Jaddoa A, Liatsis P, Hussain AJ. Uma rede neural recorrente de fechamento profundo para previsão de produção de petróleo. *Mach Learn Appl* 2021;3:100013.
- [224] Wang B, Wang J. Previsão de preços futuros de energia e modelo de avaliação com rede neural de unidade recorrente fechada bidirecional profunda e algoritmo baseado em rif. *Energia* 2021;216:119299.
- [225] Behera S, Misra R, Sillitti A. Método de prognóstico baseado em redes neurais recorrentes fechadas bidirecionais profundas multiescala para sistemas de degradação não lineares complexos. *Informe Sci* 2021;554:120–44.
- [226] Jiang P, Wang Y, Wang J. Previsão de velocidade do vento de curto prazo usando um híbrido modelo. *Energia* 2017;119:561–77.
- [227] Wang JZ, Wang Y, Jiang P. O estudo e aplicação de um novo modelo de previsão híbrido - um estudo de caso de previsão da velocidade do vento na China. *Appl Energy* 2015;143:472–88.
- [228] Fu W, Zhang K, Wang K, Wen B, Fang P, Zou F. Uma abordagem híbrida para previsão de velocidade do vento em várias etapas com base na decomposição de duas camadas, otimização híbrida DE-HHO aprimorada e KELM. *Renovar Energia* 2021;164:211–29.
- [229] Wen X. Modelagem e avaliação de desempenho de turbina eólica com base na máquina de aprendizado extremo de otimização de colônia de formigas. *Appl Soft Comput* 2020;94:106476.
- [230] Wang S, Zhang N, Wu L, Wang Y. Previsão da velocidade do vento com base na decomposição de modo empírico de conjunto híbrido e método de rede neural ga-bp. *Renovar Energia* 2016;94:629–36.
- [231] Zhang Y, Pan G, Chen B, Han J, Zhao Y, Zhang C. Modelo de previsão de velocidade do vento de curto prazo baseado em GA-ANN melhorado por VMD. *Renovar Energia* 2020;156:1373–88.

- [232] Fu W, Wang K, Li C, Tan J. Abordagem de previsão de velocidade do vento de curto prazo em várias etapas com base na análise caótica de ingredientes dominantes em várias escalas, otimização GWO-SCA híbrida aprimorada e ELM. *Energy Convers Manage* 2019;187:356–77.
- [233] Deo RC, Ghorbani MA, Samadianfard S, Maraseni T, Bilgili M, Biazar M. Modelo híbrido perceptron multicamada integrado com o algoritmo otimizador de vaga-lume para previsão da velocidade do vento do local alvo usando um conjunto limitado de dados da estação de referência vizinha. *Renovar Energia* 2018;116:309–23.
- [234] Wang J, Yang Z. Previsão de velocidade do vento de ultracurto prazo usando um algoritmo de inteligência artificial. *Renovar Energia* 2021;171:1418–35.
- [235] Yin H, Dong Z, Chen Y, Ge J, Lai LL, Vaccaro A, Meng A. Uma abordagem de decomposição secundária eficaz para previsão de energia eólica usando máquina de aprendizado extremo treinada por otimização cruzada. *Energy Convers Manage* 2017;150:108–21.
- [236] Altan A, Karasu S, Zio E. Um novo modelo híbrido para previsão de velocidade do vento combinando rede neural de memória de longo prazo, métodos de decomposição e otimizador de lobo cinzento. *Appl Soft Comput* 2021;100:106996.
- [237] Wang J, Heng J, Xiao L, Wang C. Pesquisa e aplicação de um modelo combinado baseado em otimização multi-objetivo para previsão de velocidade do vento em várias etapas. *Energia* 2017;125:591–613.
- [238] Mahmoud T, Dong Z, Ma J. Uma abordagem avançada para intervalos de previsão de geração de energia eólica ideal usando máquina de aprendizagem extrema evolutiva auto-adaptativa. *Renovar Energia* 2018;126:254–69.
- [239] Yang Z, Wang J. Uma abordagem de previsão combinada aplicada na previsão de velocidade do vento em várias etapas com base em uma estratégia de processamento de dados e um algoritmo de inteligência artificial otimizado. *Appl Energy* 2018;230:1108–25.
- [240] Chollet F, et al. Keras. 2015, <https://keras.io>.