



Análise de séries temporal usando sistema de aprendizagem de máquina baseado em metaheurística de janela deslizante para identificar padrões de consumo de energia de construção

Jui-Sheng Chou [✉], Ngoc-Tri Ngo

Departamento de Engenharia Civil e De Construção, Universidade Nacional de Ciência e Tecnologia de Taiwan, Taipei, Taiwan



h i g h l i g h t s

Este estudo desenvolve um novo sistema de previsão de janelas deslizantes de séries temporais médias.

O sistema integra metaheurística, learning de máquinas e modelos de séries temporal.

O experimento do site da infraestrutura de grade inteligente é instalado para recuperar dados em tempo real.

O sistema proposto prevê com precisão o consumo de energia em edifícios residenciais.

O sistema de previsão pode ajudar os usuários a minimizar seu uso de eletricidade.

a r t i c l e i n f o

Histórico do artigo:

Recebido em 27 de janeiro de 2016

Recebido em formulário revisado em 22 de abril de 2016

Aceito 14 maio 2016

Keywords:

Dados de grade inteligente

Gestão de energia predial

Consumo de energia

Previsão de padrões

Técnica de séries temporais

Otimização metaheurística

Aprendizado de máquina

a b s t r a c t

As redes inteligentes são uma solução promissora para a demanda de energia em rápido crescimento, porque podem aumentar consideravelmente a eficiência energética da construção. Este estudo desenvolveu um novo sistema de learning de máquina baseado em otimização metaheurística de janela de tempo para prever dados de consumo de energia de construção em tempo real coletados por uma rede inteligente. O sistema proposto integra um modelo de média móvel integrada sazonal (SARIMA) e o modelo metaheurístico de regressão vetorial support vetorial (MetaFA-LSSVR). Especificamente, o sistema proposto se encaixa no modelo SARIMA em componentes lineares de dados no primeiro estágio, e o modelo MetaFA-LSSVR captura componentes de dados não lineares no segundo estágio. Os dados em tempo real recuperados from uma grade inteligente experimental instalada em um edifício foram utilizados para avaliar a eficácia e eficácia do sistema proposto. Uma abordagem de janela deslizante k-week é proposta para empregar dados históricos como entrada para o novo system de previsão de séries temporal. O sistema de previsão rendeu altas e confiáveis taxas de precisão em previsões de 1 dia de avanço do consumo de energia predial, com uma taxa de erro total de 1,181% e erro absoluto médio de 0,026 kW h. Notavelmente, o sistema demonstra uma melhor precisão na faixa de 36,8 a 113,2% em relação aos modelos de previsão linear (ou seja, SARIMA) e modelos de previsão não linear (ou seja, LSSVR e MetaFA-LSSVR). Portanto, os usuários finais podem aplicar ainda mais as informações previstas para aumentar a eficiência do uso de energy em seus edifícios, especialmente nos horários de pico. Em particular, o sistema pode potencialmente ser ampliado para o uso da estrutura de big data para prever o consumo de energia predial.

2016 Elsevier Ltd. Todos os direitos reservados.

1. Introdução

Sistemas de rede inteligente são uma solução promissora para atender à demanda de energia em rápido crescimento [1–6]. Tais sistemas podem minimizar o impacto ambiental do consumo de energia, melhorar a energia mercados, melhorar a confiabilidade e o serviço, reduzir custos e aumentar a eficiência [7]. As redes inteligentes aplicam medidores inteligentes com

tecnologias de comunicação bidirecional para recording consumo de energia elétrica em intervalos de 1h, 15 min ou 1 min. Os medidores inteligentes então enviam as informações de volta para o centro de serviços públicos para monitoramento e faturamento. Os medidores inteligentes podem fornecer aos clientes dados de consumo de eletricidade em tempo real ou quase em tempo real. As aplicações de medidores inteligentes para alcançar uma gestão eficiente de energia predial recentemente têm atraído a atenção de vários pesquisadores [8–10].

O uso eficiente de energia predial em redes inteligentes é essencial para controlar os custos de energia, reduzir o impacto ambiental do consumo de

[✉] Autor correspondente.

Endereços de e-mail: jschou@mail.ntust.edu.tw (J.-S. Chou D10205804@mail.ntust.edu.tw) (N.-T. ONG).

energia e aumentar o valor e a competitividade do mercado dos edifícios. Prever com precisão o consumo de energia é essencial para melhorar a eficiência energética predial, pois tal previsão facilita que os proprietários de edifícios reduzam proativamente a quantidade de eletricidade consumida por eletrodomésticos, como condicionadores de ar e sistemas de iluminação. No entanto, a previsão do consumo de energia em edifícios residenciais é complexa e desafiadora das características inerentes à não linearidade e não estacionária do uso de energia em tempo real.

elétricos, como condicionadores de ar, sistemas de iluminação, aquecedores de água, micro-ondas, computadores pessoais e ventiladores elétricos. Os modelos de previsão de séries temporais são divididos em duas categorias principais, ou seja, modelos lineares e não lineares. Modelos lineares se encaixam apenas em relações lineares, enquanto modelos não lineares abordam relações altamente não lineares. Pelo que sabemos, um modelo híbrido que pode considerar adequadamente ambos os tipos de relacionamentos ainda não foi desenvolvido. Portanto, a implementação de

Nomenclatura		b0	
		r	
ARIMA		e	
SARIMA		c	
LSSVR			
ML		r _{ij}	
MetaFA		x _{i,k}	
MetaFA-LSSV			
DBMS		x _{j,k}	atratividade do vagalume a r = 0 distância entre o vagalume
SVR			de interesse e qualquer outro coeficiente de absorção
ANNs		d	constante de periculosidade
ATRÁS		yTe	distância entre quaisquer dois vagalumes i e j kth componente
AIW		hti	da coordenada espacial x _i do 1º vagalume
SERIA			kth componente da coordenada espacial x _j do jth vagalume
MAS		n	número de dimensões no espaço de pesquisa teste de consumo
RBF		x _n	de energia de dados vetor de números extraídos
R			aleatoriamente a partir de uma distribuição gaussiana ou uma
RMSE		btchaos t	distribuição uniforme no momento t rótulo numérico do valor
MAE		a0	caótico logístico individual para o nth firefly tth Gauss/mouse
MAPA		at	número caótico de iteração numérica no processo de
Maxae	autoregressive integrado móvel médio sazonal	h	otimização do parâmetro de randomização de randomização
TER	autoregressive integrado média móvel menor quadrado	tmax	inicial do parâmetro de randomização do parâmetro de
C r	suportam aprendizado de máquina de regressão vetorial	Lévy	randomização no t th geração redução de aleatoriedade
S p d	algoritmo metaheurístico de vagalume	s	constante
q P	metaheuristicfireflyalgorithm	u,v	número máximo of gerações Lévy distribuição de
D Q _{yt}	baseado no menos	y _t	energia-lei normal distribuições dados de consumo
B	quadrados suportam vetor regression	Tenente	de energia componentes lineares do consumo de
	sistema de gerenciamento de banco de	Nt	energia componentes não lineares do valor de
wq(B) hp(B)	dados suporte vetor regressão neural	^Lt	previsão de consumo de energia de SARIMA no
HP(BS)	artificial redes de algoritmo firefly	Rt	momento t residual no momento t obtido da
wQ(BS) J(x,	peso inércia adaptável	Tt	SARIMA
e) x	função de base radial	Dt Ht	temperatura ao ar livre (C) no dia da semana a hora do dia
ek xk	móvel móvel	Rle	diferença entre os dados de aprendizagem reais e previstos
yk	autoregressiva		diferença entre os dados reais e previstos de teste
N	coeficiente de correlação	HLe	aprendendo dados para temperatura ao ar livre, o dia da
ak K(x, xk)	erro quadrado médio raiz (kW h)	RTe	semana, e o hour do dia, respectivamente
b	significa erro absoluto (kW h) erro	TLe, DLe, e	Os dados de teste de HTe para temperatura ao ar livre, o dia
YLe	percentual absoluto (%) erro absoluto	TTe, DTe, e	da semana e a hora do dia, respectivamente, dados previstos
E Lep_S	máximo (kW h)		de dados de validação definem valor real do conjunto de dados
	taxa de erro total (%) parâmetro de regularização radial base	YValp	de validação
E Tep_S	núcleo período de tempo de um padrão de repetição da ordem	YVal	função objetiva do processo de otimização
	de padrão sazonal of não sazonal AR nonseasonal difere	f(C,r)	parametro mínimo e máximo de regularização
	ordem de ma não sazonal ordem de sazonal AR sazonal ordem	Cmin, Cmax	C, respectivamente
	de sazonalidade de dados sazonais da série de tempo MA para		mínimo e máximo de base radial do func-
	um polinômial de dados não estacionários em séries de tempo	rmin,	Parâmetro de tion r, respectivamente testado consumo de
	em B de graus q polinômiais em B graus p polinômias em B		energia que é previsto por SARIMA-MetaFA-LSSVR
	de graus P polinômicos em B de graus Q medidor de função de	rmax	
	otimização do aproximador linear		
	variáveis de erro padrões de entrada indicadores de amostra	YTep	
	tamanho de dados conjunto Desordeira multiplicadores viés de		
	função de kernel função de núcleo de um firefly learning		
	consumo de energia dados aprendidos consumo de energia que		
	é previsto pelo SARIMA no estágio 1 consumo de energia		
	testado que é previsto pelo SARIMA no estágio 1		

A energia consumida por um edifício residencial é considerada como dados de séries temporais que consistem em componentes lineares e não lineares. Identificar padrões de consumo de energia é difícil devido às formas altamente aleatórias de uso de energia associadas à operação de aparelhos

um modelo linear e não linear combinado inovativo é valiosa para possivelmente melhorar a previsão precisa do padrão da série temporal.

Nas últimas duas décadas, várias técnicas de previsão têm sido usadas para prever o consumo de energia de construção [11-14], e tais técnicas incluem

métodos estatísticos [15], redes neurais artificiais (ANNs) [16,17], máquinas vetorais de suporte [18], lógica difusa [19,20], modelos cinza [21,22], modelos de média móvel integrada autoregressiva (ARIMA) [23,24], e modelos sazonais de ARIMA (SARIMA) [25]. Em particular, os modelos SARIMA são um dos modelos lineares mais utilizados para previsões de séries temporais nas últimas décadas, e esses modelos podem resolver problemas de previsão relacionados à economia, engenharia e câmbio, bem como preços de ações [26]. Além disso, os modelos SARIMA têm sido amplamente utilizados para prever séries temporais sazonais, como o consumo de energia [27].

Embora os modelos SARIMA sejam razoavelmente flexíveis na previsão de séries temporais, sua maior desvantagem é que eles assumem que os dados futuros da série têm uma relação linear com dados de séries temporais anteriores. Portanto, os modelos SARIMA são inadequados para problemas não lineares do mundo real. Outras ferramentas potenciais para modelagem não linear são técnicas de aprendizado de máquina (ML), como o método de regressão vetorial de suporte menor de quadrados (LSSVR) [23]. A abordagem LSSVR reduz consideravelmente a complexidade computacional e aumenta a eficiência. A maioria dos estudos concorda que a precisão preditiva de um modelo baseado em regressão vetorial de suporte (SVR) é superior à de outros modelos baseados em ML [28,29].

O desempenho do modelo LSSVR depende muito de seus hiperparâmetros, ou seja, o parâmetro de regularização (C) e o parâmetro de função do kernel radial (r). Selecionar adequadamente os hiperparâmetros LSSVR é um problema de otimização de *challenging*. Como parte do estudo atual, um algoritmo de otimização metaheurística aprimorado foi usado para otimizar dinamicamente os hiperparâmetros LSSVR. Uma revisão da literatura indica que nenhum estudo combinou a força única de um modelo SARIMA, modelo LSSVR e um algoritmo de otimização metaheurística inspirado na natureza na previsão do consumo de energia em séries temporais.

Para preencher a lacuna de pesquisa acima mencionada, o presente estudo desenvolveu um novo sistema de aprendizado de máquina baseado em otimização metaheurística de janelas para analisar dados de séries de tempo gerados a partir de redes de rede inteligente, facilitando assim uma previsão eficiente de 1 dia de consumo de energia em edifícios residenciais. O sistema proposto integra um modelo SARIMA, um algoritmo metaheurístico metaheurístico de *firefly* inspirado na natureza (MetaFA) e o esquema LSSVR. O sistema de previsão proposto processa dados em duas etapas. Na primeira etapa, o sistema aplica o modelo SARIMA para modelar os componentes lineares do consumo de energia, enquanto no segundo estágio, aplica-se o modelo MetaFA-LSSVR para capturar os componentes não lineares do consumo de energia. O modelo MetaFA-LSSVR, que é uma combinação do esquema LSSVR e MetaFA, compensa a limitação SARIMA de abordar apenas as relações de linhaça.

A primeira originalidade deste estudo é a proposta de um sistema que combina as vantagens únicas de um método convencional de previsão de séries temporais e uma técnica de ML integrada a um algoritmo de otimização metaheurística inspirado na natureza, permitindo que o sistema capture componentes altamente lineares e não lineares dos padrões de consumo de energia; esse processo ainda não foi conduzido de forma abrangente na literatura. A segunda originalidade é o desenvolvimento de um MetaFA que pode efetivamente ajustar os hiperparâmetros LSSVR (i.e., C e r) no segundo estágio. A terceira originalidade é a proposta de uma abordagem de janela deslizante *k-week* para a uso dinamicamente de dados históricos como entrada para o novo sistema de previsão de séries temporais. Com isso, o sistema proposto pode analisar eficientemente os dados em tempo real coletados de uma infraestrutura de rede inteligente. Os usuários finais podem ainda aplicar as previsões de um dia antes para melhorar o uso eficiente de energia de aparelhos e equipamentos elétricos em seus *buildings*.

Para demonstrar a aplicabilidade do sistema de previsão, este estudo executou um experimento no local envolvendo uma rede de rede inteligente instalada em um edifício residencial. A eficácia e a confiabilidade do sistema proposto, desenvolvido no MATLAB, mostraram-se by comparando seu desempenho com

os de um modelo linear de série de referência (ou seja, SARIMA) e modelos não lineares (ou seja, LSSVR e MetaFA-LSSVR) utilizando dados em tempo real recuperados de redes de sensores e armazenados em um sistema de gerenciamento de banco de dados (DBMS). O desempenho preditivo do sistema proposto foi avaliado em termos de coeficiente de correlação (R), erro quadrado médio raiz (RMSE), erro absoluto médio (MAE), erro percentual absoluto médio (MAPE), erro absoluto máximo (MaxAE) e taxa de erro total (TER).

O restante deste artigo é organizado da seguinte forma. A Seção 2 revisa a literatura relevante, e a Seção 3 apresenta a metodologia da pesquisa. A Seção 4 descreve uma infraestrutura de rede inteligente para consumo experimental de consumo de energia em tempo real. A seção final fornece as observações finais e um esboço para trabalhos futuros.

2. Revisão da literatura

Vários pesquisadores examinaram o desempenho energético da construção [30-36]. Zhao e Magoulès revisaram abrangentemente os métodos de previsão de consumo de energia, incluindo métodos simplificados de engenharia, métodos estatísticos de séries temporais e métodos artificiais [37]. Sandels et al. usaram um método de regressão para executar uma previsão de 1 dia de antecedência do consumo de eletricidade em edifícios de escritórios [38]. O Fumo e o Fumo Bis foram aplicados análises de regressão única, múltipla e quadrática para prever o consumo diário de energia de hora em hora e diária, utilizando dados históricos de consumo de energia e parâmetros meteorológicos como insumos [15].

A previsão da série temporal é um método de previsão em valores futuros com base em dados históricos de séries temporais. Vários modelos de séries temporais, como ARIMA e SARIMA, foram desenvolvidos ao longo dos anos. Os modelos ARIMA desenvolvidos pela Box e Jenkins [23] são os *tools* estatísticos mais usados e eficazes para a previsão de séries temporais [39,40], e esses modelos constituem uma classe amplamente utilizada de modelos lineares para séries temporais univariadas. Em particular, a ARIMA combina três processos: uma função autoregressiva (AR) regredida em valores anteriores de um processo, uma função média móvel (MA) regredida em um processo puramente aleatório e uma parte integrada (I) para tornar a série de dados constante através de diferenciação.

Em uma série temporal, a sazonalidade é um padrão regular de mudanças que se repete ao longo dos períodos S, onde S é o número de períodos até que o padrão seja repetido. No entanto, a ARIMA é ineficaz na análise de dados de séries temporais caracterizadas pela sazonalidade, como o consumo de energia. A SARIMA tem sido efetivamente aplicada em inúmeros campos, como previsão de demanda de eletricidade [27], previsão de valor de produção na indústria de máquinas [41], e previsão de vendas [42].

As aplicações não lineares das técnicas de ML para prever o consumo de energia incluem o modelo ANN otimizado proposto por Li et al. para prever o consumo de energia elétrica em construção de hora em hora. Um algoritmo de otimização melhorada do enxame de partículas foi aplicado em um estudo anterior para ajustar os pesos e valores limiares de uma estrutura ANN [33]. Li et al. usaram SVR para prever a carga de resfriamento por hora de um *building* de escritório. Os resultados da simulação revelaram que a abordagem SVR é mais precisa e generalizável em comparação com o modelo tradicional de rede neural de propagação traseira [18]. Jung et al. apresentaram uma abordagem LSSVR baseada em algoritmo genético para prever o consumo de energia predial; a técnica de ML otimizada proposta é superior aos métodos convencionais [43].

Modelos híbridos lineares e não lineares têm sido empregados recentemente para prever dados de séries temporais [44]. Por exemplo, modelos híbridos compostos por modelos ARIMA e ANN têm sido amplamente utilizados pelos pesquisadores [26,45-49]. Khashei e Bijair propuseram um modelo híbrido ANN-ARIMA para previsão de séries temporais [26]. Na modelagem linear, o modelo ARIMA é benéfico para identificar e ampliar estruturas lineares existentes em dados, e o modelo ANN é usado para capturar o componente não linear dos dados, que consiste em resíduos da modelagem ARIMA. No entanto, este híbrido requer dados suficientes para fornecer um processo de modelagem eficaz. Além disso, os modelos ANN têm vários problemas que incluem a

necessidade de controlar inúmeros parâmetros, soluções incertas e a probabilidade de superequipamento.

Para superar as limitações das ANNs, um híbrido ARIMA-SVR foi usado para previsões de séries temporais, como mercado de ações [50,51] e previsão de energia elétrica [52]. Apesar de sua alta eficiência, o SVR é computacionalmente lento ao analisar grandes conjuntos de dados devido à sua velocidade que depende do número de amostras de dados [53], programação quadrática ineficiente e sua alta carga computacional para programação de otimização restrita [54]. Para reduzir a complexidade computacional, Suykens et al. propuseram o método LSSVR, uma versão aprimorada da abordagem SVR padrão. O método LSSVR resolve equações lineares em vez de um problema de programação quadrática porque altera a restrição da desigualdade em condições iguais e emprega uma função de perda quadrada [55], resultando em uma maior velocidade de cálculo e eficiência sem comprometer a vantagem do princípio da minimização de risco estrutural.

Os dois hiperparâmetros LSSVR (ou seja, C e r) desempenham papéis críticos no desempenho deste método. A sintonia fina dos hiperparâmetros supracitados é um problema de otimização; portanto, é necessário um algoritmo de otimização eficaz e eficiente. Algoritmos de otimização metaheurísticos inspirados na natureza, que podem efetivamente resolver problemas difíceis de otimização [56], são atualmente altamente comuns [57,58]. Exemplos desses algoritmos incluem o algoritmo das abelhas; otimização do enxame de partículas; pesquisa cuckoo; e o algoritmo firefly (FA). A FA é uma das mais eficientes desses algoritmos, e isso ocorre porque este algoritmo tem duas grandes vantagens sobre os outros algoritmos: capacidades de subdivisão automática e endereçamento da multimodalidade [57].

A FA tem sido efetivamente usada para resolver diversos problemas em vários domínios [59,60]. No entanto, a FA é limitada por sua tendência a ficar presa na mínima local devido ao problema da convergência prematura e a uma fraca capacidade de busca global [59]. O MetaFA desenvolvido no presente estudo supera essas deficiências incorporando a FA convencional com mapas caóticos, um peso de inércia adaptativa (AIW) e voos lévy.

Várias técnicas integradas para superar as deficiências de modelos individuais têm sido propostas por estudos anteriores. O objetivo básico dessas técnicas integradas é capitalizar as características únicas de cada modelo para capturar padrões diferentes em dados. Com base nas técnicas revisadas e seus méritos, o presente estudo propõe um híbrido dos modelos SARIMA e MetaFA-LSSVR para prever o consumo de energia.

3. Metodologia

3.1. Modelagem e previsão da série time

Em um modelo SARIMA, ar sazonal e termos médios móveis predizem dados de séries temporais usando valores de dados e erros em períodos com lags que são múltiplos de S (extensão de sazonalidade). O modelo SARIMA, denotado como SARIMA(p, d, q) (P, D, Q)_s, incorpora fatores não sazonais e sazonais em um modelo multiplicativo. Este modelo pode ser expresso como mostrado no Eq. (1), como explicado em estudos anteriores [27,61]. Eqs. o que está com o que é isso? (2)–(5) apresentar a formulação de termos em Eq. (1).

$$\begin{aligned} & \text{HPDHPDBSThed1 BSTH1 BSTH } 1/4 \text{ wq} \delta \text{BPWQ } \delta \text{BSthat} \quad \delta 1 \text{Th} \\ & \text{wq} \delta \text{BP } 1/4 \text{ } 1 \text{ w1B } \text{ w2B2 } \text{ w3B3 } \dots \text{ wqBq} \quad \delta 2 \text{This} \\ & \text{hpBTh } 1/4 \text{ } 1 \text{ h1B } \text{ h2B2 } \text{ h3B3 } \dots \text{ hpBp} \quad \delta 3 \text{This} \\ & \text{HP} \delta \text{BS } 1/4 \text{ } 1 \text{ H1} \delta \text{BS } \text{ H2} \delta \text{B2S } \text{ H3} \delta \text{B3S } \dots \text{ HP} \delta \text{BPS } \delta 4 \text{TH} \\ & \text{WQ} \quad \delta \text{BSTH } 1/4 \text{ } 1 \text{ W1} \delta \text{BS } \text{ W2} \delta \text{B2STH } \text{ W3} \delta \text{B3STH } \dots \text{ WQ } \delta \text{BQSTH5TH} \end{aligned}$$

onde p representa a ordem AR não sazonal, d representa diferenciação não sazonal, q representa a ordem móvel não sazonal (MA), P representa a ordem ar sazonal, D representa diferenciação sazonal, Q representa a ordem de MA sazonal, S representa o período de tempo de um padrão sazonal repetitivo, e B representa o operador de turno para trás para um item de dados de série temporal não estacionário y_t . Além disso, $wq(B)$, $hp(B)$, $HP(BS)$ e $WQ(BS)$ são polinômios em B de graus q , p , P e Q , respectivamente, onde $wq(B)$ e $WQ(BS)$ indicam que y_t é uma função do erro de previsão anterior na previsão de y_t , e $hp(B)$ e $HP(BS)$ indicam que y_t é uma função de seus próprios valores anteriores; at é uma interferência atual. Normalmente, at é considerada o residual estimado no momento em que t . Além disso, p , d , q , P , D e Q são todos inteiros, e $(1-B)^d y_t$ pode ser convertido em uma série estacionária usando o operador de diferença $1-B$; B satisfaz

$$BYT = y_{tk} \text{ e } Bkyt = y_{tk}.$$

Detalhes de um modelo SARIMA também são fornecidos em [23]. Em particular, ciclos iterativos de quatro etapas são necessários para se encaixar em um modelo SARIMA [27,61], e esses ciclos são descritos da seguinte forma: (1) Identifique a estrutura do SARIMA(p, d, q) (P, D, Q)_s modelo, (2) estimar parâmetros desconhecidos, (3) realizar testes de bondade de ajuste nos resíduos estimados e (4) prever resultados futuros com base nos dados conhecidos. Como mencionado, o modelo SARIMA é amplamente utilizado para prever dados sazonais de séries temporais devido à sua vantagem de eliminar o efeito da periodicidade no processo de previsão [27].

Inúmeras aplicações acadêmicas e industriais do modelo SARIMA foram desenvolvidas nas últimas três décadas [41]. Embora o modelo SARIMA seja amplamente utilizado para análise da demanda de eletricidade e seja uma abordagem de alta precisão para a previsão de dados sazonais, erros no processo de previsão são inevitáveis [27]. Notavelmente, a maior limitação deste modelo é sua suposição de que os valores da série temporal têm uma estrutura de correlação linear. Especificamente, o modelo SARIMA não pode capturar adequadamente padrões não lineares em estruturas de dados complexas, como o consumo de energia. Portanto, no presente estudo, uma técnica de ML baseada em otimização metaheurística inspirada na natureza foi integrada ao modelo SARIMA para evitar a limitação acima mencionada deste modelo.

3.2. Otimização metaheurística inspirada na natureza na técnica de aprendizado de máquina

3.2.1. Aprendizado de máquina

As técnicas de ML provaram-se recentemente viáveis e eficazes na análise de dados de séries temporais [41,46,50,62-64]. A abordagem SVR desenvolvida pela Vapnik em 1995 [65] é uma técnica de ML baseada na teoria da aprendizagem estatística e no princípio da minimização de riscos estrutural. Apesar de sua alta eficiência, a abordagem SVR é computacionalmente lenta ao analisar grandes conjuntos de dados porque sua velocidade depende do número de amostras de dados e solucionadores de programação quadrática [53]. Para aumentar a velocidade computacional, Suykens et al. [55] propôs o método LSSVR. Esta técnica ML possui muitos recursos avançados que permitem alta capacidade de generalização e computação rápida.

O processo de treinamento LSSVR implica o uso de uma função de custo de menos quadrados para obter um conjunto linear de equações em um espaço duplo para minimizar o custo computacional. Assim, métodos iterativos como o método de gradiente conjugado são tipicamente usados para derivar uma solução, resolvendo eficientemente um conjunto de equações lineares [63]. Dado um conjunto de dados de treinamento fxk ; $y_{kNk1/41}$, a estimativa de função usando LSSVR é formulada como um problema de otimização, conforme expresso no Eq. (6):

$$\min J_{\delta x}; eTH \quad \frac{1}{2} \sum_{t=1}^N 4kxk^2_{-t} \quad 12CXe2k;$$

$x; b; e$

$k/41$

sujeito a $y_k 1/4 h x; u d x k t h t h e k; k 1/4 1;::: N \delta 6 T H$

onde $J(x, e)$ denota a função de otimização, x denota o parâmetro aproximador linear, e_k $2 R$ denotar variáveis de erro, $C P 0$ denota uma constante de regularização especificando a constante representação do trade-off entre erro empírico e flatness função, x_k denota padrões de entrada, y_k denota rótulos de previsão, e N denota o tamanho da amostra.

O modelo LSSVR resultante para estimativa de função é expresso como Eq. (7).

N

$f \delta x T h 1/4 x a k K \delta x; x k t h 7 t h$

$k/41$

onde a_k ; b denotar os multiplicadores Lagrange e o termo de viés, respectivamente, e $K(x, x_k)$ denota a função do kernel. No espaço de recurso, a função do kernel pode ser expressa como Eq. (8):

m

$K \delta x; x k T h 1/4 X g k k x T h a t x T h i s$

$k/41$

A função radial gaussiana (RBF) e a função polinomial são funções de kernel comumente utilizadas. O presente estudo utilizou a função do kernel RBF por duas razões: (1) Tem menor complexidade matemática em comparação com a função do kernel polinomial, e (2) resolve efetivamente problemas altamente não-claros, como a previsão de consumo de energia. A função RBF pode ser expressa matematicamente como mostrado no Eq. (9):

$K \delta x; x k T h 1/4 \exp d k x x k k 2 = 2 r 2 T h i s$

onde r é a largura RBF.

O modelo LSSVR resolve efetivamente problemas de previsão; no entanto, sua principal desvantagem é que sua precisão depende dos hiperparâmetros LSSVR. Portanto, selecionar hiperparâmetros apropriados é essencialmente um problema de otimização. A otimização do hiperparâmetro no modelo LSSVR deve incluir C e r . Portanto, como parte das contribuições deste estudo, um algoritmo de otimização metaheurística inspirado na natureza (ou seja, MetaFA) é proposto para otimizar os hiperparâmetros LSSVR C e r .

3.2.2. Algoritmo de otimização metaheurística inspirado na natureza

A FA, desenvolvida por Yang [66], é um algoritmo metaheurístico estocástico inspirado na natureza que identifica simultaneamente e efetivamente o optima global e o optima local de um problema de otimização. A FA convencional aplica três regras idealizadas [66]: (1) Cada vagalume é atraído por outros vagalumes, independentemente do sexo, porque todos os vagalumes são unissex; (2) a atratividade de um vagalume é proporcional ao seu brilho e diminui com a distância aumentante entre este e outros vagalumes, e um vagalume também se move aleatoriamente se nenhum outro vagalume for mais brilhante; e (3) o brilho de um vagalume é afetado ou determinado pelo espaço de busca da função objetiva.

Para um problema de maximização, o valor de brilho é simplesmente definido como uma proporção do valor da função objetiva. Como a atratividade de um vagalume é proporcional à intensidade da luz visível aos vagalumes adjacentes, a atratividade b pode ser calculada usando Eq. (10).

onde b é a atratividade de um vagalume, b_0 é a atratividade do vagalume em $r = 0$, r é a distância entre o vagalume de interesse e qualquer outro vagalume, e e é um coeficiente constante, e c é o coeficiente de absorção.

A distância entre os vagalumes i e j em x_i e x_j , respectivamente, é a distância cartesiana:

$$r_{ij} = \sqrt{\sum_{k=1}^D (x_{i,k} - x_{j,k})^2} \quad (11)$$

onde r_{ij} é a distância entre quaisquer vagalumes i e j em x_i e x_j , respectivamente; $x_{i,k}$ é o componente k th da coordenada espacial x_i da décima sócia; $x_{j,k}$ é o componente k th da coordenada espacial x_j do j th firefly; e D é o número de dimensões no espaço de busca.

Eq. (12) especifica o movimento do décimo primeiro vagalume quando atraído pelo j th firefly que é mais atraente (mais brilhante):

$$x_{ti+1} = x_{ti} + b_0 e^{c r_{ij}} (x_{tj} - x_{ti}) \quad (12)$$

onde x_{ti} é a coordenada do primeiro vagalume na $(t + 1)$ th iteração; x_{ti} é a coordenada do décimo primeiro vagalume na iteração t ; x_{tj} é a coordenada do j th firefly na iteração t ; c é o coeficiente de absorção, que normalmente varia de 0,1 a 10 na maioria das aplicações; a é um parâmetro de randomização; e h_{ti} é um vetor de números extraídos aleatoriamente de uma distribuição gaussiana ou uma distribuição uniforme no momento t .

Em geral, os três parâmetros são c , b e a . Quando $b_0 = 0$, o movimento do vagalume é uma simples caminhada aleatória. Notavelmente, os termos de randomização na FA se estenderam facilmente a outras distribuições, como mapas caóticos e voos lévy. Os resultados de análise de sensibilidade ideal de c são obtidos quando $c = 1$; $b_0 = b_{min}$ = a atratividade em $r_{ij} = 0$.

Apesar de sua eficiência em inúmeras aplicações, a FA tende a ficar presa em um ideal local [59]. Definir os parâmetros para melhorar a convergência é outro desafio encontrado na aplicação deste algoritmo. Os parâmetros de controle da FA devem ser otimizados para equilibrar a exploração e a exploração [67]. Portanto, o MetaFA incorpora três componentes metaheurísticos, ou seja, mapas caóticos, os voos AIW e Lévy, na FA convencional para melhorar seus recursos de busca e otimização.

Primeiro, na implementação da FA, um mapa Gauss/mouse é usado para sintonizar um parâmetro de attractiveness e um mapa logístico é usado para fornecer uma população inicial altamente diversificada. Em segundo lugar, a AIW é ajustada dinamicamente no processo de otimização para controlar as capacidades de exploração locais e globais da FA. Em terceiro lugar, os voos da Lévy aceleram a busca local gerando novos bairros ideais em torno da solução ideal derivada. Fig. 1 apresenta o pseudocódigo do MetaFA. A subseção a seguir define cada termo e elucida o papel de cada componente.

3.2.2.1. Gerando diversidade da população inicial por meio do mapa logístico. Para melhorar a diversidade de soluções iniciais e a qualidade da população inicial, um mapa logístico é utilizado para gerar uma população inicial altamente diversificada na fase inicial do MetaFA. A fórmula do mapa logístico é apresentada da seguinte forma:

$$x_{n+1} = 1/4 g x_n 1 - x_n \quad (13)$$

3.3. Sistema de aprendizado de máquina baseado em otimização metaheurística de janela deslizante da série time

Capturar padrões de consumo de energia em edifícios residenciais é difícil porque esses padrões dependem muito do comportamento dos proprietários de edifícios ao longo do tempo. Uma estratégia híbrida que exhibe habilidades de modelagem linear e não linear é uma alternativa adequada para prever o consumo de energia. Os modelos SARIMA e MetaFA-LSSVR capturam características de dados em componentes lineares e não lineares, respectivamente. Como a abordagem híbrida proposta combina os pontos fortes de cada modelo, ele captura efetivamente padrões de dados de séries temporais. O consumo de energia é considerado um function de componentes lineares e não lineares, como descrito no Eq. (22):

$$C_t = Y_t + \sum_{i=1}^n L_i + \sum_{j=1}^m N_j \quad (22)$$

onde Y_t denota o consumo de energia, L_i denota os componentes lineares, e N_j denota os componentes não lineares.

Ambos os tipos de componentes devem ser determinados para prever o uso de energia future. O sistema de aprendizagem de máquina baseado em otimização de janelas deslizantes de séries temporais (ou seja, SARIMAMetaFA-LSSVR) compreende duas etapas, ou seja, estágio 1, envolvendo modelagem linear univariada, e estágio 2, envolvendo modelagem não linear multivariada. O primeiro e o segundo estágios implicam o uso dos modelos SARIMA e MetaFA-LSSVR para modelar os componentes lineares e não lineares dos dados de consumo de energia, respectivamente. O princípio principal do sistema de predição proposto é que o modelo SARIMA primeiro modela os componentes lineares de acordo com Eq. (23):

$$\hat{Y}_t = \sum_{i=1}^p \phi_i Y_{t-i} + \sum_{j=1}^q \theta_j \epsilon_{t-j} + \epsilon_t \quad (23)$$

onde \hat{Y}_t é o valor de previsão do modelo SARIMA no momento t e ϵ_t é o residual no momento t como obtido a partir deste modelo.

Os resultados da primeira etapa (ou seja, valores de previsão e resíduos de modelagem linear) são utilizados na segunda etapa. Os valores de previsão e os resíduos, bem como a temperatura ao ar livre, tipo de dia (ou seja, dia da semana ou fim de semana), e hora do dia (ou seja, 0, 1, ..., 23) são utilizados como entradas para o modelo MetaFA-LSSVR. Assim

$$N_j = \sum_{i=1}^p \phi_i \epsilon_{t-i} + \sum_{j=1}^q \theta_j \epsilon_{t-j} + \epsilon_t \quad (24)$$

onde ϵ_t é a temperatura ao ar livre na hora t , D_t é o dia da semana na hora t (ou seja, segunda, terça-feira, quarta-feira, quinta, sexta, sábado e domingo), e H_t é a hora do dia (ou seja, 0, 1, 2, ..., 23). Portanto, a previsão combinada é apresentada como Eq. (25).

$$Y_t = \sum_{i=1}^p \phi_i \hat{Y}_{t-i} + \sum_{j=1}^q \theta_j \epsilon_{t-j} + \epsilon_t \quad (25)$$

onde lag é o tempo de atraso para a previsão. O valor da defasagem pode ser determinado realizando uma análise de sensibilidade.

Fig. 2 ilustra o procedimento envolvido na construção do sistema de aprendizagem de máquina metaheurística de janela deslizante de série temporal proposto (SARIMA-MetaFA-LSSVR). No estágio 1 (modelagem linear univariada), os dados históricos de consumo de energia são recuperados espontaneamente a partir de um DBMS de nuvem de dados de rede inteligente em tempo real. O DBMS é um componente de infraestrutura de medição de rede inteligente em tempo real instalado no edifício residencial examinado neste estudo (Seção 4.1). Os dados recuperados são divididos automaticamente em dados históricos de aprendizagem da k -week (Y_{Le}) e no décimo dia de $(k+1)$ -week datas de teste (Y_{Te}) ($i = 1-7$ para segunda a domingo).

Os dados de aprendizagem são usados para treinar e construir um modelo linear de previsão de séries temporais, SARIMA, enquanto os dados de teste são usados para avaliar o modelo SARIMA após o processo de aprendizagem.

As saídas da fase 1 incluem o consumo energy aprendido ($Y_{Le,S}$) e o consumo de energia testado ($Y_{Te,S}$), que são previstos pelo modelo SARIMA com base nos conjuntos de dados de aprendizagem e teste, respectivamente. Além disso, o R_{Le} residual é a diferença entre os dados reais e de aprendizagem previstos do consumo de energia, e o R_{Te} é a diferença entre os dados reais e previstos de teste de consumo de energia.

Na fase 2 (modelagem multivariada), os dados de saída (ou seja, Y_{Le} , Y_{Te} , $Y_{Le,S}$, $Y_{Te,S}$, R_{Le} e R_{Te}), bem como a temperatura ao ar livre, tipo de dia e hora do dia recuperado da nuvem DBMS no estágio 1 servem como entradas para previsão não linear. Nesta etapa, o modelo MetaFA-LSSVR é usado para modelar os componentes não lineares dos dados de consumo de energia. O MetaFA sintoniza automaticamente os hiperparâmetros LSSVR (i.e., C e γ). A otimização dos hiperparâmetros LSSVR C e γ minimiza erros de previsão, melhorando assim o desempenho preditivo geral.

Para otimização, o modelo MetaFA-LSSVR continua usando atributos de dados históricos de learning da k -week (ou seja, Y_{Le} , $Y_{Le,S}$, R_{Le} , T_{Le} , D_{Le} e H_{Le}) e no décimo dia de $(k+1)$ -semana atributos de dados de teste (ou seja, Y_{Te} , $Y_{Te,S}$, R_{Te} , T_{Te} , D_{Te} e H_{Te}). O $P\%$ ($0 < P < 100$) de dados de aprendizagem (ou seja, dados de treinamento) é usado para treinar o modelo MetaFA-LSSVR, enquanto os $(100 - P)\%$ dos dados de aprendizagem (ou seja, dados de validação) são usados para validar os hiperparâmetros LSSVR ideais. O valor P pode ser personalizado através de experimentos de teste para um determinado conjunto de dados e está definido para 60-90 na maioria dos casos; no presente estudo, P foi definido como 70. Posteriormente, os dados de teste são usados para avaliar o desempenho de aprendizagem do modelo LSSVR otimizado.

O valor RMSE dos dados de validação é usado como uma função objetiva padrão para otimizar o processo de seleção do hiperparâmetro. (26) descreve a função objetiva do MetaFA:

$$\text{RMSE}_{\text{OptimizationValidation}} = \frac{1}{n} \sum_{i=1}^n \sqrt{\frac{1}{2} (\hat{Y}_i - Y_i)^2} \quad (26)$$

il/41

Espaço de busca: C_{\min} a C_{\max} e γ_{\min} a γ_{\max} . Número máximo de iterações: t_{\max} .

where $f(C, \gamma)$ is the objective function of the MetaFA

RMSEOptimizationValidação-dados é o RMSE calculado usando os dados de validação previstos (Y_{Valp}) e valor de validação real (Y_{Val}), n é o tamanho da amostra de dados usada para validação, e C_{\min} , C_{\max} , γ_{\min} , γ_{\max} e t_{\max} são constantes definidas pelo usuário com base em uma série de experimentos de teste. C e γ são otimizadas quando o critério de parada é atingido.

Depois que os hiperparâmetros LSSVR são obtidos de acordo com o processo de otimização acima mencionado, o modelo LSSVR ideal é avaliado ainda mais usando os dados de teste. Combinando com os componentes lineares das saídas do estágio 1, as medidas previstas de consumo de energia (Y_{Te}) e de previsão de desempenho são determinadas na fase 2. O ótimo sistema SARIMA-MetaFA-LSSVR prevê o futuro consumo de energia um dia antes com um intervalo de 15 minutos. Todos os resultados de análise são enviados automaticamente de volta ao DBMS em nuvem e ao servidor web para visualização de padrões.

3.4. Método de avaliação de desempenho

Este estudo utilizou medidas de R , RMSE, MAE, MAPE, MaxAE e TER para avaliar a precisão preditiva do sistema proposto. Eqs. (27)–(32) mostram as respectivas fórmulas utilizadas para o cálculo dessas medidas:

1/4 nmaxX yyi
iyi il/41
MaxAE 1/4th 0jThis

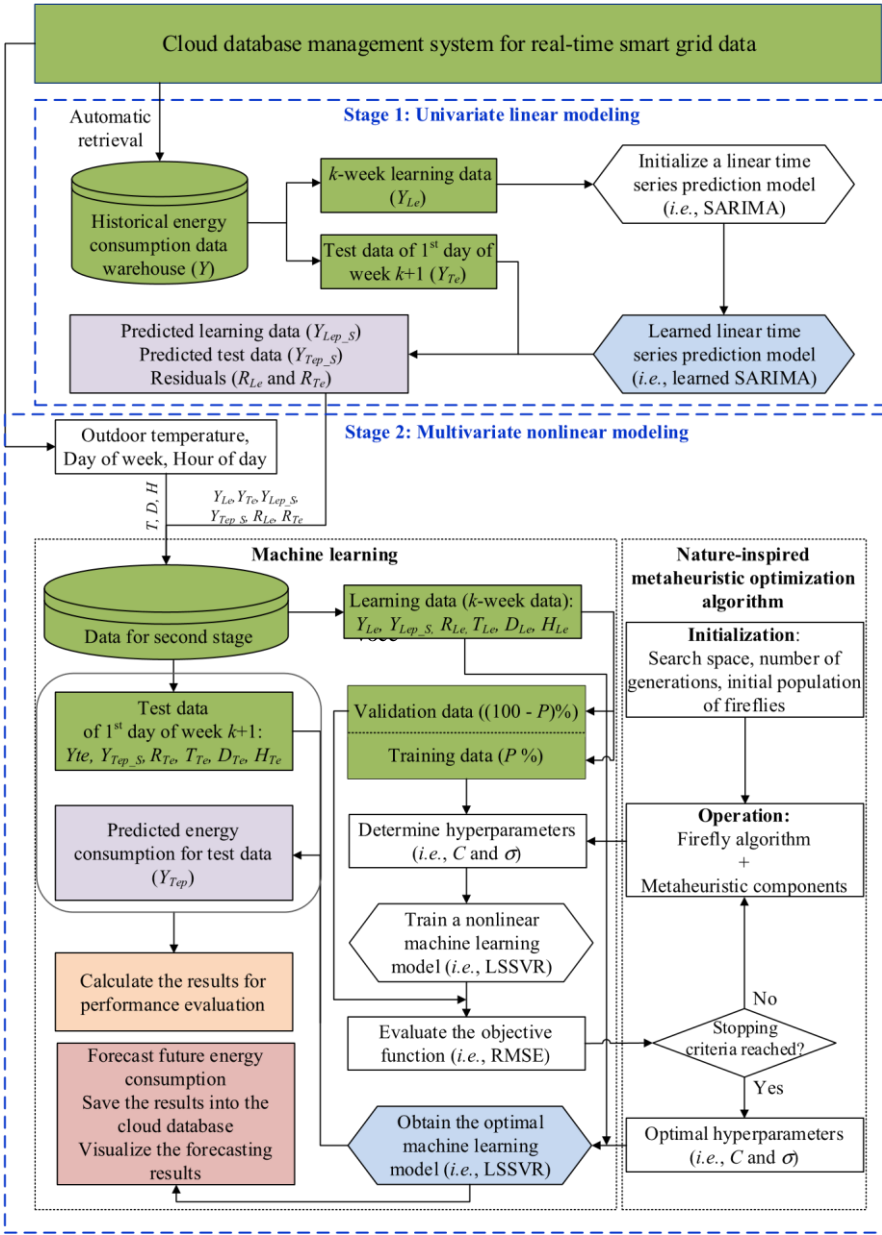


Fig. 2. Sistema de aprendizagem de máquina baseado em otimização metaheurística baseada em otimização de janelas de séries temporal.

il/41

você

$$\frac{1}{4} qn P_{y2i}^P y_i - 2ffiqPn\delta PPy0i2THPpn\delta PPy0i2$$

$$R_i = \frac{n}{y} \frac{y}{y_0}$$

$$RMSE \frac{1}{4} ut1_Xn y0iP2ffn\delta28Que$$

il/41

$$MAE = \frac{1}{N} \sum_{i=1}^N \frac{X}{1/4jyi y0ij\delta29This}$$

$$\frac{MAPE1 nyi y0i30}{TER \frac{1}{4} jPni141Py0i - mny ni1/41yij\delta32P}$$

1/4 onde y0 representa o valor previsto, y representa o valor real, e n representa o número de amostras de dados.

4. Experimento no site e recuperação de dados

4.1. Infraestrutura de medição de rede inteligente em tempo real

Os dados de consumo de energia em tempo real coletados de uma rede de rede inteligente foram necessários para avaliar a aplicabilidade do sistema de predição proposto na identificação de padrões de consumo de energia predial. Assim, uma infraestrutura de medição de grade inteligente foi instalada em um

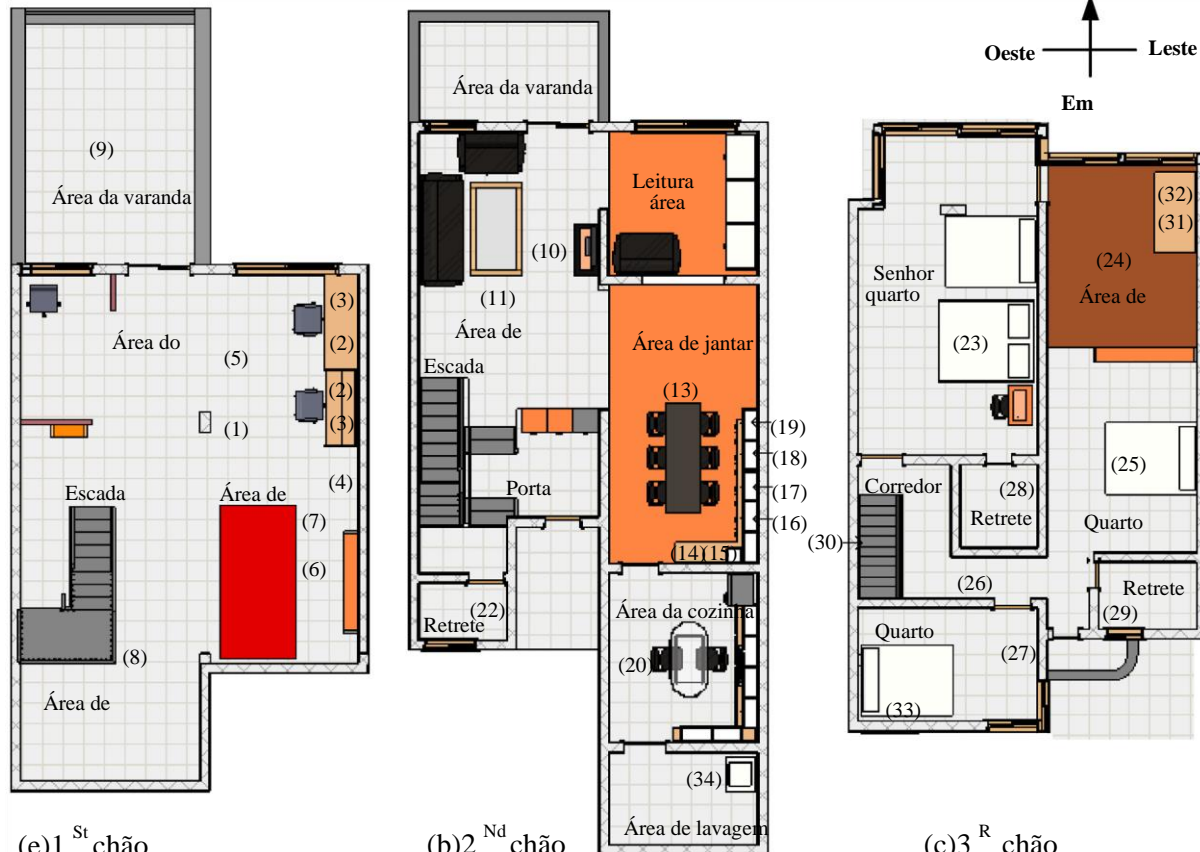


Fig. 3. Layout de construção e localização de aparelhos. (a) 1st chão, (b) 2nd chão, (c) 3rd chão. Resumo do nome e quantidade dos aparelhos e equipamentos elétricos utilizados em cada andar do edifício. Suas posições são mapeadas para os pisos correspondentes na Fig. 3, utilizando o número atribuído dos aparelhos e equipamentos na coluna "Localização" da Tabela 1.

A infraestrutura de medição de rede inteligente rastreou o uso de energia de aparelhos e dispositivos elétricos no edifício em tempo real. Fig. 4 ilustra os componentes da infraestrutura de medição de rede inteligente em tempo real, indicando que ela compreende uma infraestrutura de medição e subseção, rede de comunicação e dados

Fig. 3. Layout de construção e localização de aparelhos.

edifício residencial experimental para monitorar dados de sensores em tempo real. O experimento foi ambientado em um típico edifício de três andares localizado no distrito de Xindian, Cidade de Nova Taipei, Taiwan. Este prédio era ocupado por uma família de cinco membros composta por três crianças e seus pais. Fig. 3 apresenta o layout do prédio

e posições de aparelhos e equipamentos elétricos, indicando que o edifício experimental tem uma área total de piso de 350 m².

O primeiro andar consistia principalmente em uma área de escritório/teste, onde o proprietário administrava seu negócio, enquanto o segundo andar era composto principalmente por uma cozinha, além de refeições, leitura e áreas de convivência; o terceiro andar continha principalmente um quarto principal, quartos infantis, sala de hóspedes e sala de estudo. A Tabela 1 apresenta um infraestrutura de gestão. O componente de medição e subseção de infraestrutura compreende medidores inteligentes e submeters com recursos bidirecionais de comunicação que permitem a coleta, a captura e o controle das informações. A rede de comunicação transfere as informações coletadas por cada dispositivo de medição para um DBMS. Detalhes sobre as configurações da infraestrutura de grade inteligente são fornecidos em [69]. Assim, as análises de grade inteligente podem ser implementadas para fornecer aos usuários feedback, como o uso previsto de energia.

O DBMS (ou seja, MySQL) armazena dados em tempo real recuperados da infraestrutura de medição de rede inteligente, incluindo dados de consumo de energia, informações sobre o aparelho, dados de sensores de temperatura e umidade e resultados analíticos, como previsão de eletricidade futura e alternativas de economia de eletricidade. Além disso, o DBMS armazena parâmetros elétricos como tensão, corrente, potência, frequência e fator de potência. Os dados de consumo de eletricidade são recuperados dos medidores inteligentes e transferidos para um servidor dedicado através da rede de comunicação. O fluxo de dados dos medidores inteligentes chega a intervalos de 1 min, resultando em 1440 pontos de dados por dia de cada medidor inteligente.

4.2. Descrição e pré-processamento de dados

Considerando o ciclo de cálculo da demanda de energia elétrica pelas companhias elétricas, os dados foram convertidos de intervalos de 1 min em intervalos de 15 minutos por meio da execução de um procedimento de conversão no MySQL. Consequentemente, 96 pontos de dados para a construção de consumo de energia foram armazenados diariamente. Além disso, temperatura ao ar livre, dia da semana (ou seja, segunda, terça, quarta, quinta, sexta, sábado e domingo), e hora do dia (ou seja, 0, 1, 2, ..., 23) foram coletados neste banco de dados. Fig. 5 mostra um 15-min

Tabela 1
Construção de eletrodomésticos e equipamentos elétricos.

No.	Nome do equipamento	Quantidade	Área	Localização em Fig. 3
1º andar				
1	Ventilador elétrico	2	Área do escritório	(1)
2	Computador pessoal	2	Área do escritório	(2)
	Lâmpada 3Desk	2	Área do escritório	(3)
Iluminação	4Wall	1	Área do escritório	(4)
Iluminação	5Cell 1	1	Área do escritório	(5)
Iluminação	6Cell 2	1	Área de teste	(6)
7	Máquinas especializadas	1	Área de teste	(7)
	8Dehumidificador	1	Área de teste	(8)
Iluminação	9Cell 3		Varanda	(9)
2º andar				
	10TV definido	1	Área de convivência	(10)
	11Cell iluminação 1	1	Área de convivência	(11)

12	Ventilador elétrico 1	1	Área de convivência	(12)
Iluminação	13Cell 2	1	Área de jantar	(13)
	14Oven	1	Área de jantar	(14)
	15Microwave	1	Área de jantar	(15)
	16Rice cozinheiro	1	Área de jantar	(16)
Máquina	de água 17Hot	1	Área de jantar	(17)
	Aquecedor de água 18Hot	1	Área de jantar	(18)
	19Stereo	1	Área de jantar	(19)
Iluminação	20Cell 3	1	Área da cozinha	(20)
	21Refrigerator	1	Área da cozinha	(21)
22	Máquina de lavar	1	Área da cozinha	(34)
Iluminação	23Small	1	Retrete	(22)
3° andar				
Iluminação	23Cell	4		(23) (24) (25) (26)
Iluminação	24Wall	4		(27) (28) (29) (30)
25	Computador pessoal	1	Área de estudo	(31)
	26Lamp	1	Área de estudo	(32)
Condicionamento	27Air	1	Quarto	(33)

MySQL Query Browser - Connection: jasonlin@140.118.205.145:3306 / ntust

File Edit View Query Script Tools Window Help

Transaction Explain Compare

Resultset 1

SQL Query Area

1 SELECT * FROM `ntust`.`input`

Datetime_index	Temp...	Weekday	Hour	Total_Energy_...	2F_Lighting_...	2F_Living_area	1F_Office_area	3F_Energy_Co...	3F_Air_Condition
2015-06-24 00:00:00	28.2	2	0	0.587	0.1475	0.114675	0.026775	0.0275	0.197
2015-06-24 00:15:00	28.2	2	0	0.545	0.146	0.114325	0.026625	0.01375	0.01725
2015-06-24 00:30:00	27.5	2	0	0.442	0.01625	0.0071	0.027075	0.0295	0.19725
2015-06-24 00:45:00	27.3	2	0	0.385	0.016	0.006725	0.026375	0.0155	0.0175
2015-06-24 01:00:00	28.1	2	1	0.313	0.016	0.00665	0.026775	0.0155	0.191
2015-06-24 01:15:00	27.1	2	1	0.367	0.016	0.0067	0.02645	0.0155	0.0175
2015-06-24 01:30:00	27.5	2	1	0.313	0.01575	0.00665	0.026525	0.0155	0.01775
2015-06-24 01:45:00	28.3	2	1	0.302	0.016	0.0068	0.02645	0.0155	0.193
2015-06-24 02:00:00	27.6	2	2	0.3	0.01575	0.006775	0.026425	0.01525	0.1875
2015-06-24 02:15:00	27.8	2	2	0.308	0.016	0.006775	0.0265	0.015	0.0175
2015-06-24 02:30:00	27.4	2	2	0.306	0.01575	0.00675	0.0288	0.0145	0.01775
2015-06-24 02:45:00	28.2	2	2	0.296	0.01575	0.0067	0.02645	0.01425	0.01775
2015-06-24 03:00:00	28.3	2	3	0.306	0.016	0.0068	0.026375	0.01425	0.018
2015-06-24 03:15:00	27.7	2	3	0.288	0.016	0.0068	0.02655	0.01425	0.01775
2015-06-24 03:30:00	28.1	2	3	0.315	0.01575	0.00675	0.026475	0.01425	0.01775
2015-06-24 03:45:00	28.1	2	3	0.281	0.01575	0.006825	0.02705	0.01425	0.01775
2015-06-24 04:00:00	27.9	2	4	0.305	0.01575	0.0068	0.027275	0.01425	0.01775
2015-06-24 04:15:00	27.8	2	4	0.276	0.016	0.006775	0.026325	0.01425	0.018
2015-06-24 04:30:00	27.9	2	4	0.301	0.01575	0.006775	0.026275	0.0145	0.1435
2015-06-24 04:45:00	28	2	4	0.28	0.016	0.00675	0.026325	0.01425	0.18825
2015-06-24 05:00:00	27.5	2	5	0.305	0.01575	0.006775	0.026275	0.01425	0.186
2015-06-24 05:15:00	27.1	2	5	0.286	0.016	0.006825	0.026575	0.01425	0.19
2015-06-24 05:30:00	27.8	2	5	0.305	0.016	0.00675	0.026625	0.0145	0.018
2015-06-24 05:45:00	28	2	5	0.281	0.016	0.006775	0.026375	0.0145	0.018

2879 rows fetched in 0.3946s (0.0997s)

Edit Apply Changes Discard Changes First Last Search

Figo. 5Coleta de dados em tempo real de Quinze ata Não banco de dados MySQL.

Tabela 2

Atribuas informações usadas para o sistema de predição proposto.

Atributo	Descrição
Datetime_index	Momento em que os dados são gravados
Temperatura	Temperatura ao ar livre (°C)
Semana	Dia da semana, que tem valores de 0 = Segunda- feira, 1 = Terça-feira, 2 = Quarta- feira, 3 = Quinta-feira, 4 = Sexta-feira, 5 = Sábado, 6 = Domingo
Hora	Hora do dia, que tem valores de 0, 1, 2, ... 21, 22 e 23
Total_Energy_Consumption	Consumo total de energia de construção (kW h)
2F_Lighting_Energy_Consumption	Consumo de energia do sistema de iluminação no segundo andar (kW h)
2F_Living_area	Consumo de energia da área de convivência no segundo andar (kW h)
1F_Office_area	Consumo de energia da área de escritório no primeiro andar (kW h)

o padrão de consumo de energia tem uma característica sazonal diária

(baixo durante o dia e alto à noite). Seu padrão flutua dra...

3F_Energy_Consumption	Consumo de energia de eletrodomésticos no banheiro, quarto e quarto de hóspedes no terceiro andar (kW h)	maticamente por causa dos comportamentos de uso de energia amplamente variados dos moradores do edifício.
3F_Air_Condition	Consumo de energia de um ar condicionado no terceiro andar (kW h)	A eficácia do sistema de predição proposto foi avaliada por meio de dados experimentais registrados pela infraestrutura de rede inteligente instalada no edifício

residencial. Especificamente, a eficácia do sistema foi avaliada utilizando-se o consumo total de energia predial durante as 5 semanas de 22 de junho de 2015 a 26 de julho de 2015. Os dados de consumo de energia utilizados para a avaliação foram o consumo total de energia predial de todos os aparelhos e equipamentos elétricos apresentados na [Tabela 1](#). O desempenho de previsão do sistema proposto foi avaliado dividindo os dados definidos em conjuntos de aprendizado e teste. Além disso, o conjunto de dados foi ajustado diariamente com a chegada de novos dados para formar uma nova janela deslizante.

Para validação, o processo de avaliação foi repetido diariamente durante uma semana, resultando em sete avaliações. [Fig. 7](#) mostra a abordagem da janela deslizante k-week para análise de séries temporais, em que o sistema de previsão usou $k = 4$ semanas de dados históricos para aprendizado e, em seguida, previu o consumo de energia do edifício residencial para o dia seguinte.

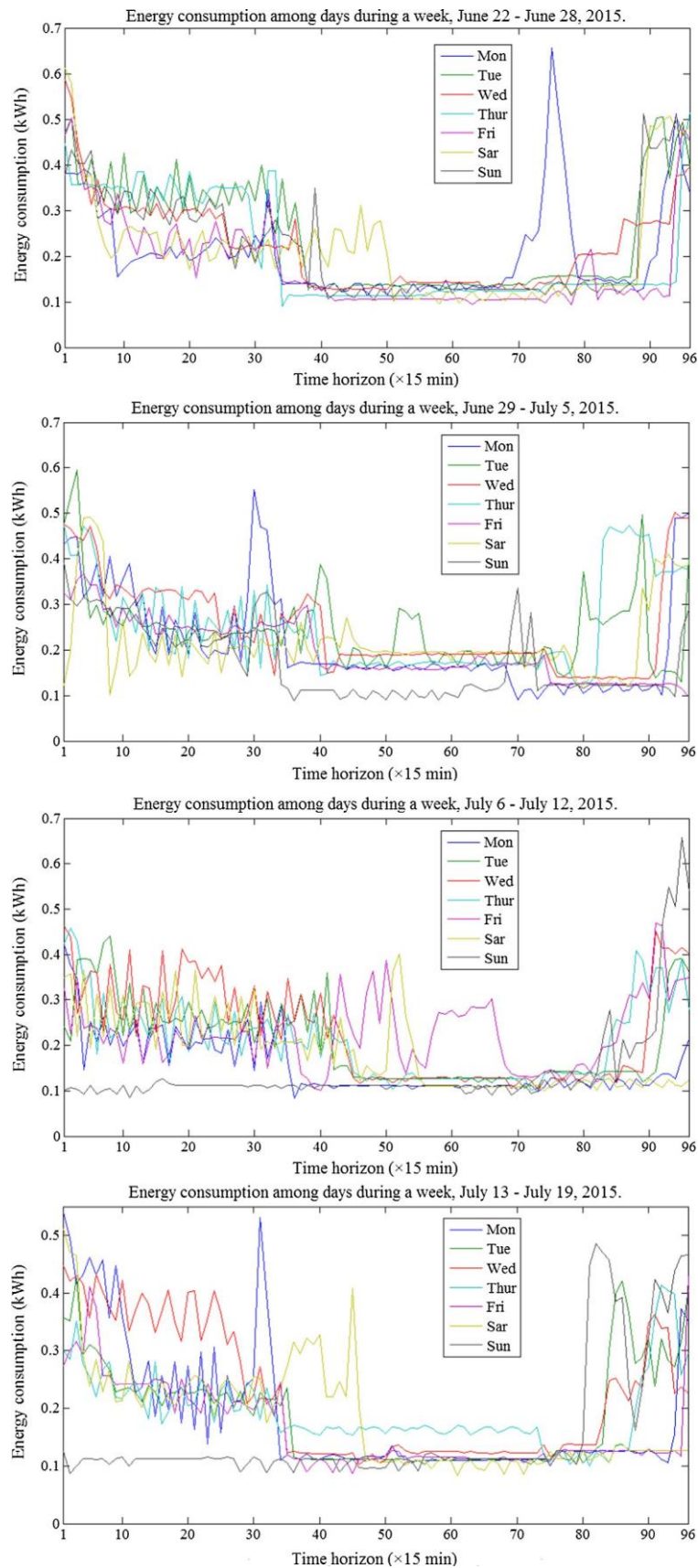
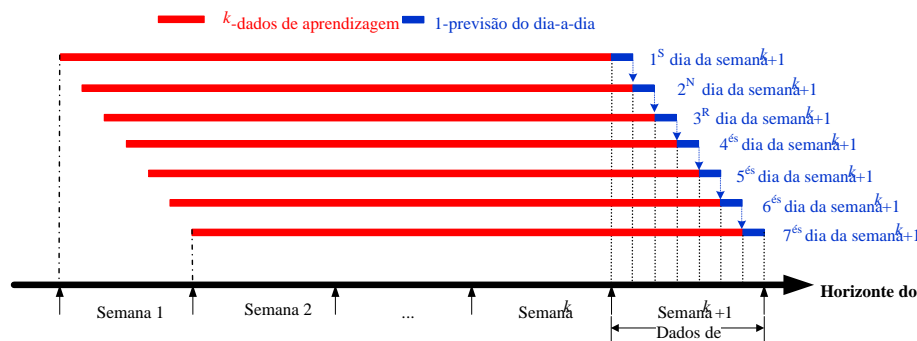


Fig. 6. O consumo real de energia está entre os dias da semana por 4 semanas.



Figo. 7E k- abordagem janela Deslizante semana durante análise de séries

Tabela 3

Configurações de parâmetros para o sistema de previsão proposto.

Parâmetro	Descrição	Ambiente
Primeira etapa: Modelo SARIMA p		
d	Ordem autoregressiva não sazonal	1
q	Diferença não sazonal	0
P	Média móvel não sazonal	1
D	Ordem autoregressiva sazonal	96
Q	Diferença sazonal	0
S	Ordem média móvel sazonal	96
	Tempo de repetição do padrão sazonal	96
Segunda etapa: Modelo MetaFA-LSSVR		
Faixa de C de r	Parâmetro de regularização do LSSVR	[1031012]
Não. de vagalumes	Função de kernel sigma de RBF no tamanho populacional LSSVR do MetaFA	[¹⁰³¹⁰³]
Critérios de parada	Geração máxima	50
	Três taxas consecutivas de mudança nos valores da função objetiva	25
		<108
Mapa caótico logístico	Gerar população inicial com alta diversidade	Geração aleatória baseada em Eq. (13)
b _{min}	Valor mínimo do parâmetro de atratividade b	0.1
c	Coefficiente de absorção	1
Mapa caótico gauss/mouse	Ajuste automaticamente o parâmetro b	Parâmetro atraente baseado em Eq. (15)
um	Aleatoriedade do movimento do vagalume	a ₀ = 0.2
AIW	Controle as capacidades de exploração locais e globais do algoritmo do enxame	h = 0.9
Voo Levy	Acelere a busca local gerando novas soluções em torno da solução ideal	s = 1,5

5. Análise e discussão de dados

5.1. Previsões do sistema

A Tabela 3 apresenta um resumo das configurações experimentais utilizadas para o sistema de previsão SARIMA-MetaFA-LSSVR proposto. Para o modelo SARIMA, foi realizada uma análise de cenário para otimizar as configurações do sistema de pré-candidatura. As configurações obtidas através da análise do cenário foram $p = 1$, $d = 0$, $q = 1$, $P = 96$, $D = 0$, $Q = 96$ e $S = 96$. A Tabela 3 também mostra um resumo das configurações do modelo MetaFA-LSSVR, a segunda etapa do system de previsão proposto. C e r foram definidos para faixas de [0,001, 1012] e [0,001, 103], respectivamente. O tamanho populacional do MetaFA (ou seja, hiperparmetros LSSVR de C e r) foi definido para 50. Os hiperparmetros LSSVR foram considerados otimizados quando o criterion parando foi atingido (ou seja, após 25 iterações ou 3 taxas consecutivas de mudança nos valores da função objetiva são inferiores a 108).

5.2. Análise de Sensibilidade

Conforme mencionado na Seção 3.3, o tempo de atraso da previsão, lag, deve ser identificado antes da previsão. O objetivo de determinar esse tempo de atraso é identificar a defasagem ideal para a previsão. Nesta análise, os insumos foram dados de consumo de energia, e os valores de defasagem foram definidos para variar de 1 a 96 (Não. de pontos de dados por dia) para comparar o

desempenho de previsão; a faixa de valores de defasagem foi utilizada devido às características sazonais dos padrões de consumo de energia (ou seja, $S = 96$). A análise foi realizada diariamente durante 7 dias, e a Tabela 4 apresenta um resumo dos resultados de comparação de desempenho. Os resultados analíticos confirmaram que o sistema de previsão obteve o maior desempenho com um valor de defasagem de 6.

A segunda análise foi um exame paramétrico para identificar os insumos aproximados para previsão. Quatro cenários foram comparados: Cenário 1 considerado apenas consumo de energia histórico como insumo; cenário 2 considerado temperatura externa como insumo, além de dados históricos de consumo de energia; cenário 3 considerado dados históricos de consumo de energia, temperatura ao ar livre, tipo de dia (ou seja, dia ou fim de semana), e hora do dia (ou seja, 0, 1, ..., 23); e cenário 4 considerado consumo histórico de energia, tipo de dia (ou seja, dia da semana ou fim de semana), e hora do dia como insumos.

As atividades diárias de construção de moradores normalmente diferem entre weekdays e finais de semana e entre o dia e a noite; por exemplo, os moradores podem trabalhar durante o dia e voltar para casa à noite. Como esses comportamentos afetam o uso de aparelhos, influenciam o consumo de energia. A análise de sensibilidade 2 foi realizada uma vez por dia durante uma semana por meio do sistema proposto. A Tabela 5 lista as medidas médias de desempenho, ou seja, RMSE, MAE, MAPE, R e TER, obtidas para os quatro cenários utilizando os dados do teste. Os resultados de comparação confirmaram que o cenário 4 rendeu resultados superior aos dos outros cenários. Portanto, dados históricos de consumo de energia, tipo de dia e hora do

Tabela 4
Comparações de desempenho na análise de sensibilidade 1.

Foi	Medida média de desempenho em dados de teste					Classi	Foi	Medida média de desempenho em dados de teste					Classi
	RMSE (kW h)	MAE (kW h)	MAPA (%)	R	TER (%)			RMSE (kW h)	MAE (kW h)	MAPA (%)	R	TER (%)	
	10.179	0.032	18.97	0.736	5.40	5	47	0.282	0.083	51.99	0.421	15.93	96
	20.178	0.032	18.45	0.725	5.13	2	48	0.223	0.051	30.81	0.590	3.21	37
	30.177	0.032	18.33	0.718	5.00	3	49	0.282	0.087	50.89	0.462	5.61	95
	40.179	0.032	18.34	0.718	4.64	4	50	0.246	0.065	39.78	0.544	4.62	85
	50.180	0.033	18.51	0.680	3.38	7	51	0.238	0.058	35.36	0.546	4.38	74
	60.175	0.031	18.01	0.729	4.10	1	52	0.233	0.055	33.93	0.531	7.68	81
	70.183	0.034	18.93	0.662	4.25	11	53	0.237	0.057	34.99	0.541	5.20	77
	80.181	0.034	18.58	0.675	3.61	8	54	0.227	0.053	34.85	0.526	6.89	76
	90.186	0.036	19.45	0.674	3.13	9	55	0.264	0.074	45.59	0.491	5.41	92
	100.191	0.038	20.23	0.656	2.96	13	56	0.234	0.055	34.70	0.551	5.21	69
	110.189	0.037	20.28	0.647	3.72	14	57	0.222	0.050	32.04	0.597	8.99	63
	120.188	0.036	20.09	0.654	3.39	12	58	0.225	0.051	31.75	0.584	5.51	46
	130.199	0.041	22.98	0.639	2.87	19	59	0.227	0.052	32.31	0.586	6.84	60
	140.182	0.034	19.04	0.716	4.08	6	60	0.229	0.053	32.95	0.567	6.14	65
	150.190	0.036	21.11	0.663	5.80	16	61	0.258	0.071	43.29	0.502	6.59	91
	160.188	0.036	19.98	0.683	3.61	10	62	0.238	0.060	38.19	0.576	10.65	88
	170.193	0.039	21.51	0.661	3.09	15	63	0.227	0.052	33.14	0.572	5.17	56
	180.199	0.041	23.17	0.648	3.05	18	64	0.265	0.074	46.98	0.515	9.74	94
	190.207	0.044	25.01	0.629	4.76	25	65	0.235	0.058	37.28	0.577	7.11	80
	200.199	0.041	24.06	0.651	4.25	21	66	0.226	0.052	33.02	0.609	7.51	58
	210.210	0.046	25.43	0.624	3.93	26	67	0.234	0.056	35.85	0.598	7.78	75
	220.206	0.044	25.14	0.635	4.10	22	68	0.225	0.051	32.13	0.601	7.81	57
	230.200	0.041	23.11	0.646	3.46	20	69	0.237	0.057	37.57	0.573	6.66	78
	240.198	0.041	22.88	0.650	3.03	17	70	0.224	0.051	32.21	0.591	7.10	55
	250.207	0.044	25.28	0.600	3.39	23	71	0.221	0.049	30.91	0.598	7.43	47
	260.210	0.045	26.84	0.581	6.18	33	72	0.222	0.050	32.20	0.595	9.71	66
	270.215	0.047	27.77	0.583	5.78	35	73	0.222	0.050	31.63	0.580	6.83	52
	280.209	0.046	27.65	0.578	8.01	40	74	0.232	0.056	35.23	0.566	8.48	82
	290.218	0.049	27.63	0.599	3.49	31	75	0.255	0.068	44.39	0.523	10.53	93
	300.219	0.049	28.98	0.552	5.29	42	76	0.229	0.053	34.17	0.585	8.39	71
	310.229	0.053	30.85	0.562	4.91	53	77	0.253	0.066	42.77	0.544	7.57	49
	320.218	0.049	29.42	0.588	4.49	34	78	0.226	0.052	33.17	0.583	6.71	39
	330.210	0.045	26.62	0.648	4.35	24	79	0.232	0.055	34.45	0.565	9.13	41
	340.210	0.044	26.46	0.608	4.34	28	80	0.248	0.064	41.32	0.521	7.95	30
	350.213	0.047	27.24	0.607	3.59	29	81	0.224	0.052	32.35	0.595	8.32	51
	360.227	0.052	32.12	0.543	9.97	79	87	0.232	0.055	34.86	0.564	9.41	84
	370.209	0.045	26.73	0.622	3.90	27	88	0.209	0.044	27.37	0.615	7.77	32
	380.225	0.051	31.80	0.603	5.69	45	89	0.226	0.052	32.22	0.585	6.96	59
	390.227	0.053	31.61	0.553	4.22	49	90	0.211	0.045	28.07	0.630	8.59	36
	400.221	0.049	29.90	0.596	5.04	39	91	0.216	0.047	29.68	0.638	9.76	43
	410.224	0.051	30.36	0.576	4.19	41	92	0.223	0.051	30.93	0.584	7.29	54
	420.215	0.048	28.15	0.599	3.80	30	93	0.228	0.054	32.15	0.559	6.78	68
	430.226	0.054	32.43	0.584	5.22	51	94	0.224	0.052	31.18	0.599	8.48	61
	440.239	0.058	36.58	0.527	8.11	86	95	0.229	0.054	32.07	0.568	6.92	67
	450.235	0.056	34.52	0.539	4.90	73	96	0.223	0.051	31.00	0.597	7.25	50
	460.236	0.057	34.34	0.547	4.98	72							

Nota: Valores em negrito indicam maior desempenho da medida correspondente em comparação com outros.

Tabela 5
Comparação de cenários na análise de sensibilidade 2.

RMSEMAEMAPE1991.5733.43 (kW h) (kW h) (%)266752273.26460.32						
1	0.178	0.181	0.033	0.033	18.34	18.98
2	0.179	0.033	18.88	0.723	5.18	4
3	0.172	0.030	17.38	0.718	4.36	3
4				0.743	4.41	1
Cenário	Medida média de desempenho		em dados de teste		Classi	
			R	TER (%)		

Nota: Valores em negrito indicam maior desempenho da medida correspondente em comparação com outros.

Tabela 6

Hiperparâmetros LSSVR otimizados.		
Avaliação	Parâmetro de regularização (C)	Parâmetro de função do kernel (r)
3	288176792.72	542.43
4	140795309.83 94987405.73	573.84
5	31624073.58 994.45	445.86
6		1000.00
7		22.06

dia foram usados como insumos para prever o uso de energia predial neste estudo.

5.3. Resultados de avaliação e discussão

Este estudo aplicou um processo de validação envolvendo sete avaliações para minimizar qualquer viés no sistema de predição proposto. Para demonstrar a eficácia e confiabilidade do proposto

Tabela 7

Desempenho do sistema proposto nas fases de aprendizagem e teste.

Dado	Medir	Ter.				Sex	Sáb.	Chão	De.	Max	Min
Meu.						Casar.	Thur.				
Conjunto de dados de aprendizagem	R	0.911	0.918	0.917	0.915	0.915	0.905	0.920	0.914	0.920	0.905
	RMSE (kW h)	0.161	0.158	0.158	0.159	0.159	0.163	0.155	0.159	0.163	0.155
	MAE (kW h)	0.026	0.025	0.025	0.025	0.025	0.026	0.024	0.025	0.026	0.024
	MAPA (%)	12.808	12.406	12.504	12.797	12.797	13.351	12.372	12.719	13.351	12.372
	MaxAE (kW h)	0.346	0.294	0.352	0.358	0.358	0.324	0.311	0.335	0.358	0.311
	TER (%)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Conjunto de dados de teste	R	0.845	0.831	0.933	0.833	0.833	0.836	0.482	0.799	0.933	0.482
	RMSE (kW h)	0.167	0.189	0.168	0.168	0.168	0.160	0.105	0.161	0.189	0.105
	MAE (kW h)	0.028	0.036	0.028	0.028	0.028	0.026	0.011	0.026	0.036	0.011
	MAPA (%)	14.599	15.467	13.731	15.942	15.942	17.000	10.817	14.785	17.000	10.817
	MaxAE (kW h)	0.197	0.186	0.200	0.252	0.252	0.176	0.034	0.185	0.252	0.034
	TER (%)	1.641	2.728	1.768	4.556	4.556	2.733	1.212	2.742	4.556	1.212
Tempo de CPU decorrido		1.179	1.133	1.125	1.345	1.35	1.0368	1.104	1.181	1.345	1.037

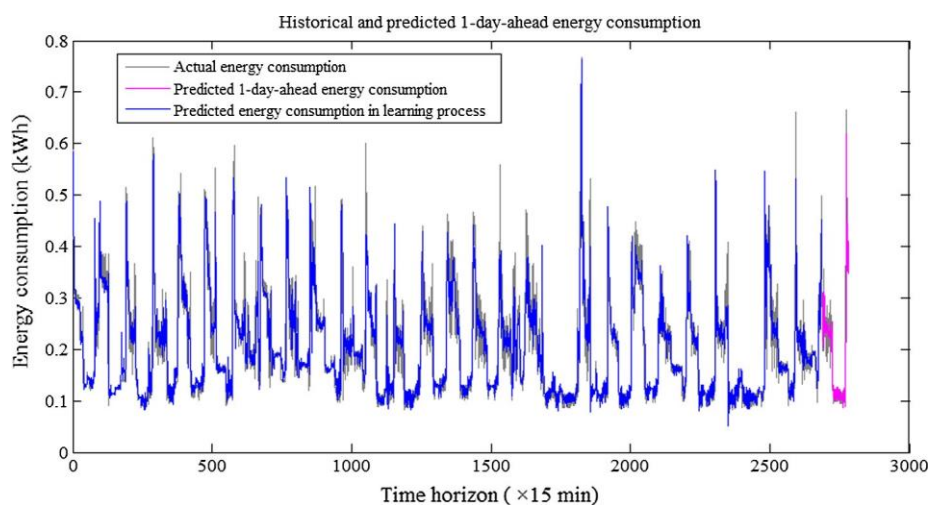


Fig. 8. Consumo real e previsto de energia obtido utilizando dados de aprendizagem e teste.

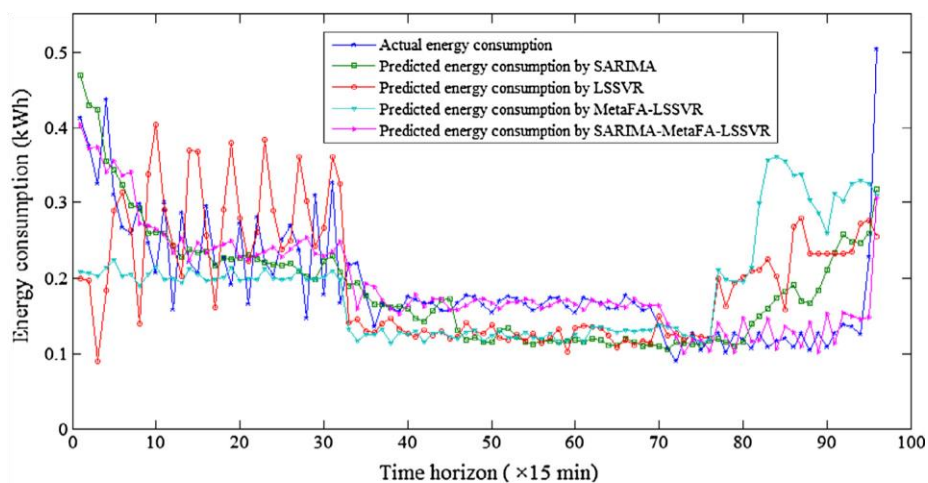
Tabela 8

Comparações de desempenho entre o sistema proposto e o modelo SARIMA na fase de teste.

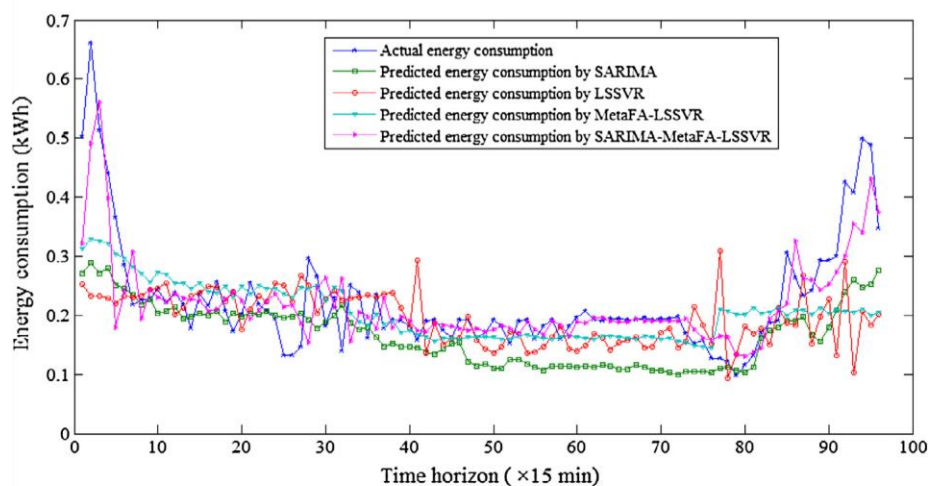
Dados	Medir	Meu.	Ter.	Casar.	Thur.	Sex.	Sáb.	Sol.	Av.	Max	Min
SARIMA	R	0.749	0.723	0.506	0.506	0.847	0.520	0.038	0.556	0.847	0.038
	RMSE (kW h)	0.054	0.090	0.105	0.105	0.047	0.069	0.082	0.079	0.105	0.047
	MAE (kW h)	0.044	0.066	0.058	0.058	0.034	0.049	0.065	0.053	0.066	0.034

LSSVR	MAPA (%)	25.224	27.050	23.319	23.319	21.052	37.451	65.429	31.835	65.429	21.052
	MaxAE (kW h)	0.185	0.372	0.474	0.474	0.150	0.232	0.204	0.299	0.474	0.150
	TER (%)	1.433	25.874	9.247	9.247	5.348	18.397	62.518	18.866	62.518	5.348
	Tempo de CPU (s)	0.181	0.180	0.225	0.222	0.193	0.170	0.169	0.191	0.225	0.169
	R	0.304	0.226	0.457	0.561	0.538	0.364	0.173	0.375	0.561	0.173
	RMSE (kW h)	0.270	0.253	0.284	0.222	0.235	0.260	0.318	0.263	0.318	0.222
MetaFA-LSSVR	MAE (kW h)	0.073	0.064	0.081	0.049	0.055	0.068	0.101	0.070	0.101	0.049
	MAPA (%)	41.264	26.282	46.263	34.092	38.564	55.834	99.551	48.836	99.551	34.092
	MaxAE (kW h)	0.253	0.428	0.454	0.228	0.248	0.212	0.482	0.329	0.482	0.212
	TER (%)	3.680	12.918	1.654	13.934	2.837	31.300	98.565	23.556	98.565	1.654
	Tempo de CPU (s)	1.088	1.131	1.143	1.138	1.198	1.146	1.173	1.145	1.198	1.088
	R	0.077	0.538	0.159	0.574	0.711	0.589	0.322	0.424	0.711	0.159
SARIMA-MetaFA-LSSVR (sistema proposto)	RMSE (kW h)	0.274	0.235	0.318	0.237	0.215	0.250	0.251	0.254	0.318	0.215
	MAE (kW h)	0.075	0.055	0.101	0.056	0.046	0.063	0.063	0.066	0.101	0.046
	MAPA (%)	46.099	23.198	54.708	39.838	31.078	49.837	60.561	43.617	60.561	31.078
	MaxAE (kW h)	0.249	0.332	0.983	0.265	0.178	0.204	0.213	0.346	0.983	0.178
	TER (%)	1.805	8.570	8.311	21.024	1.369	33.038	54.826	18.421	54.826	1.369
	Tempo de CPU (s)	0.965	1.113	1.124	1.094	1.111	1.433	2.189	1.290	2.189	0.965
	R	0.845	0.831	0.933	0.833	0.833	0.836	0.482	0.799	0.933	0.482
	RMSE (kW h)	0.167	0.189	0.168	0.168	0.168	0.160	0.105	0.161	0.189	0.105
	MAE (kW h)	0.028	0.036	0.028	0.028	0.028	0.026	0.011	0.026	0.036	0.011
	MAPA (%)	14.599	15.467	13.731	15.942	15.942	17.000	10.817	14.785	17.000	10.817
	MaxAE (kW h)	0.197	0.186	0.200	0.252	0.252	0.176	0.034	0.185	0.252	0.034
	TER (%)	1.641	2.728	1.768	4.556	4.556	2.733	1.212	2.742	4.556	1.212
	Tempo de CPU (s)	1.179	1.133	1.125	1.345	1.35	1.0368	1.104	1.181	1.345	1.037

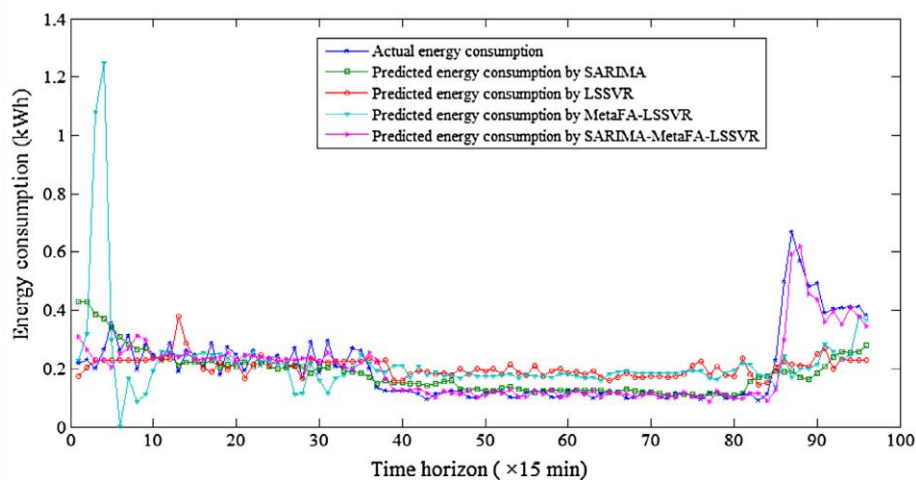
Nota: Valores ousados denotam as melhores medidas de desempenho entre os modelos.



(a) Avaliação 1 na segunda-feira.



(b) Avaliação 2 na terça-feira.



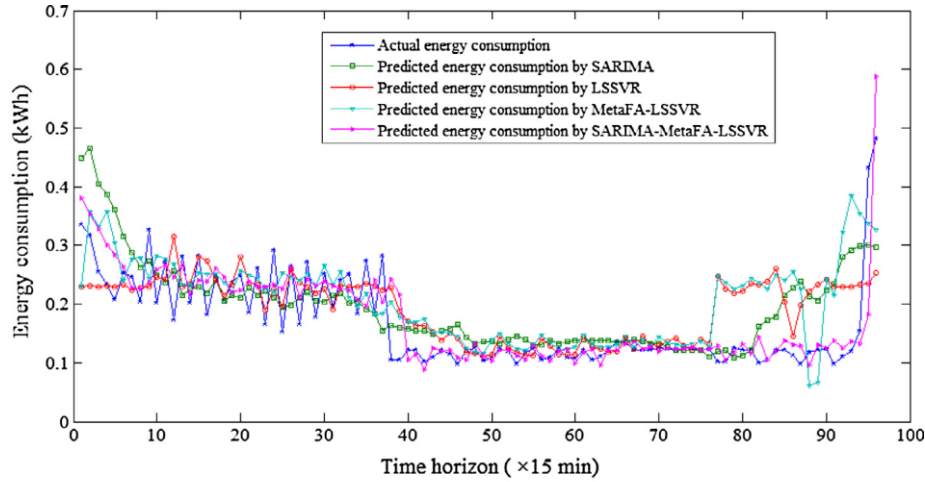
(c) Avaliação 3 na quarta-feira.

Fig. 9. Consumo real e previsto de energia por 1 semana obtidos com dados de teste.

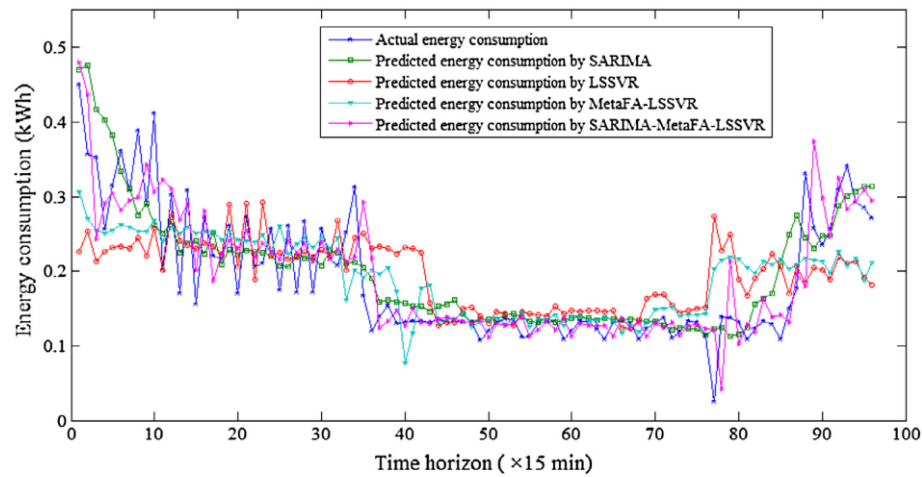
pôr. De acordo com os resultados de otimização, as faixas de C e r para previsão do uso de energia do edifício residencial considerado neste estudo foram $[991,57, 2,88E+8]$ e $[22,06, 1000,00]$, respectivamente, no procedimento de validação de rolamento.

A Tabela 7 mostra um resumo do desempenho preditivo do sistema proposto nas fases de aprendizagem e teste envolvendo sete avaliações, uma em cada dia da semana. Os valores médios de R, RMSE, MAE, MAPE, MaxAE e TER derivados dos dados de consumo de energia foram de 0,799, 0,161 kW h, 0,026 kW h, 14,785%, 0,18 kW h e 2,742%, respectivamente, demonstrando a eficácia do sistema. O menor TER derivado quando os dados de domingo foram utilizados foi de 1,212%. Além disso, o tempo médio de computação da CPU foi

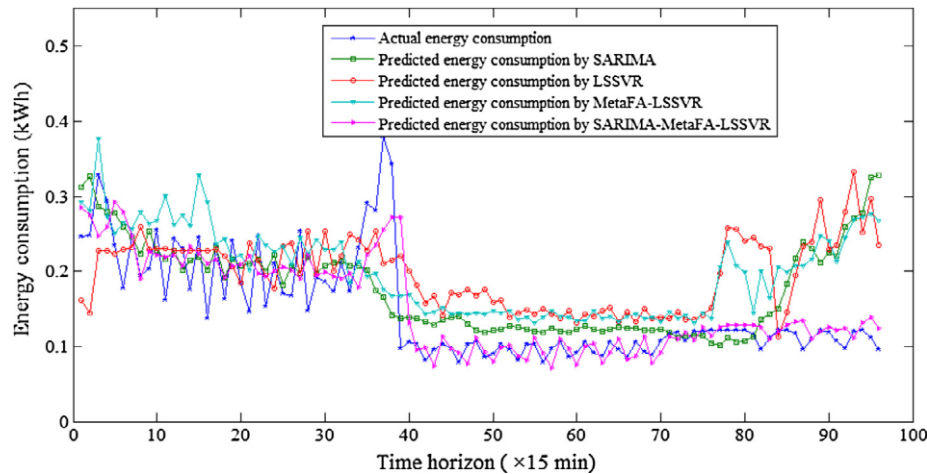
de apenas 1.181 s, o que é eficiente. Fig. 8 mostra um enredo do consumo real de energia e previsão de consumo de energia no processo de aprendizagem (linha azul) e no processo de teste (linha rosa), significando que o sistema proposto capturou com precisão o uso de energia predial, mesmo quando oscilou amplamente devido aos padrões complexos de uso de equipamentos elétricos pelos moradores do edifício.



(d) Avaliação 4 na quinta-feira.

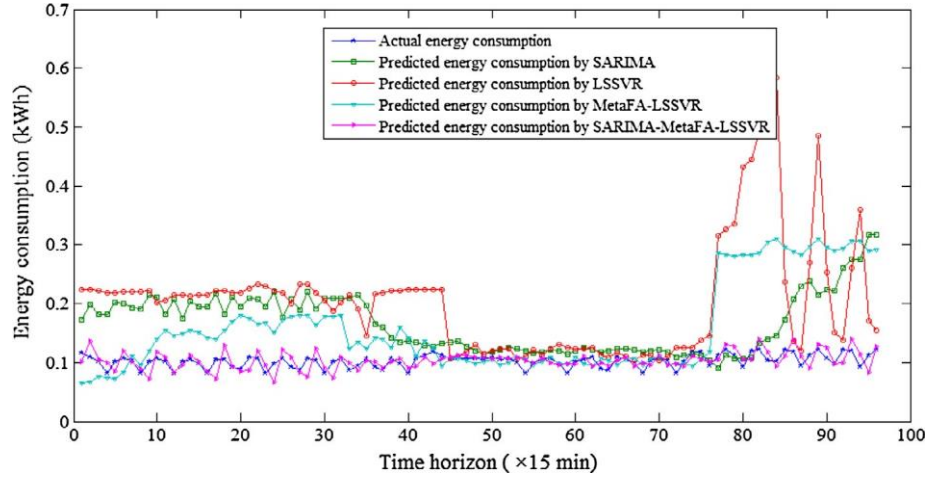


(e) Avaliação 5 na sexta-feira.



(f) Avaliação 6 no sábado.

O desempenho preditivo do sistema SARIMA-MetaFA-LSSVR foi comparado com os do SARIMA (modelo linear de linha de base), LSSVR (modelo não linear da linha de base) e MetaFA-LSSVR



g Avaliação 7 no domingo.

Fig. 9 (continuado)

Tabela 9
Melhoria percentual alcançada pelo sistema proposto em relação aos outros modelos.

Modelo/método	Medição de desempenho					Melhoria pelo modelo proposto (%)				
	R	RMSE (kW h)	MAE (kW h)	MAPA (%)	TER (%)	R	RMSE	MAE	MAPA	TER
SARIMA	0.556	0.079	0.053	31.835	18.866	43,8*	—	50,5*	53,6*	85,5*
LSSVR	0.375	0.263	0.070	48.836	23.556	113,2*	38,9*	62,3*	69,7*	88,4*
MetaFA-LSSVR	0.424	0.254	0.066	43.617	18.421	88,3*	36,8*	59,7*	66,1*	85,1*
SARIMA-METAFA-LSSVR	0.799	0.161	0.026	14.785	2.742					

Nota: As medições de melhorias de desempenho e resultados de testes de hipóteses foram calculadas utilizando-se medidas médias de desempenho. Valores ousados denotam as melhores medidas de desempenho entre os modelos. * Indica que um nível de significância superior a 1%.

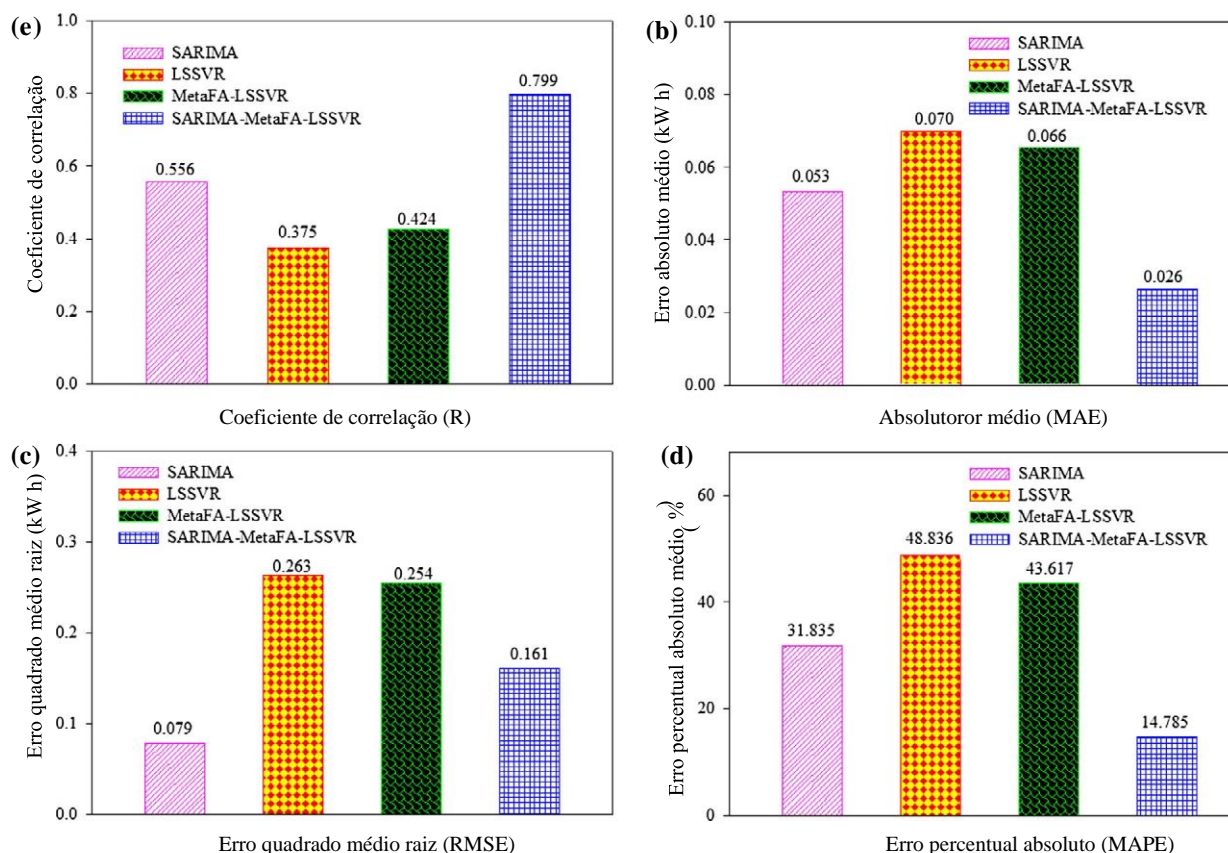


Fig. 10. Comparações de desempenho entre o sistema proposto e os outros modelos.

(modelo não linear otimizado) modelos para demonstrar a superioridade do sistema proposto. A Tabela 8 apresenta a comparação de desempenho utilizando os dados do teste para cada dia da semana. Os resultados nesta tabela confirmam que o sistema proposto supera os modelos other, conforme indicado pelas medidas de desempenho. Especificamente, a média do TER do sistema proposto (2,742%) foi superior às dos modelos SARIMA (18,866%), LSSVR (23,556%) e MetaFA-LSSVR (18,421%) Fig. 9 fornece uma comparação do sistema proposed e dos outros modelos de acordo com o consumo real e previsto de energia obtido utilizando dados diários de teste por 7 dias. O número revela que as previsões do sistema proposto estão mais próximas dos padrões reais, em comparação com as de outros modelos.

A Tabela 9 mostra a melhoria de precisão percentual alcançada pelo sistema proposto em comparação com os modelos SARIMA, LSSVR e MetaFA-LSSVR. Em comparação com os outros modelos, o sistema proposto apresentou valores R, RMSE, MAE, MAPE e TER significativamente melhorados na faixa de 36,8 a 113,2% em um nível de significância de 1%. Portanto, o sistema proposto pode prever com precisão padrões de consumo de energia em edifícios residenciais. A comparação das medidas de desempenho (Fig. 10) confirma a superioridade do sistema proposto em relação aos outros modelos.

O sistema de aprendizagem de máquina baseado em otimização metaheurística de janela de tempo proposto serve como uma ferramenta de previsão facilitando os proprietários de edifícios no monitoramento e previsão da consumação de energia em tempo real. Os moradores podem reduzir proativamente o consumo de energia de eletrodomésticos e equipamentos elétricos com base nos resultados de previsão em tempo real do sistema proposto. Trabalhar no desenvolvimento de um sistema de suporte de decisão baseado na Web para otimizar a alocação de recursos energéticos para fornecer aos usuários uma ferramenta fácil e conveniente para os usuários está atualmente em andamento. Neste sistema baseado na Web, os cronogramas

operacionais dos aparelhos também podem ser otimizados com antecedência, permitindo que os residentes tomem facilmente planos de ação alternativa para economizar o uso de energia ligando ou desligando os aparelhos em um determinado período. O sistema proposto pode ser potencialmente ampliado na gestão de big data relacionados ao consumo de energia em rede inteligente.

6. Conclusões

Este estudo propõe um novo sistema de previsão baseado em otimização metaheurística para analisar dados de séries temporais gerados a partir de uma rede inteligente para prever eficientemente o consumo de energia com um dia de antecedência. O sistema de predição proposto integra o modelo MetaFA-LSSVR (um modelo não linear otimizado) com o modelo SARIMA (um modelo linear) e pode compensar as limitações do modelo SARIMA. Além disso, o sistema proposto explora a técnica de série temporal e a otimização metaheurística no aprendizado de máquina. Agents metaheurísticos (mapas caóticos, voos Lévy e AIW) são incorporados à FA padrão para efetivamente ajustar os hiperparâmetros LSSVR.

A eficácia e a eficiência do sistema proposto foram demonstradas pela análise dos dados em tempo real coletados de uma infraestrutura de medição de rede inteligente. Especificamente, usando a abordagem da janela deslizante, este estudo utilizou dados em tempo real recuperados de uma rede de rede inteligente para avaliar a eficácia do sistema proposto. Uma análise de sensibilidade confirma que um tamanho de janela de 4 semanas e um valor de defasagem ideal de 6 foram adequados para analisar dados históricos para obter previsões de 1 dia antes. Além disso, dados históricos de consumo de energia, tipo de dia e hora do dia foram determinados como atributos críticos para prever padrões de uso de energia.

O sistema SARMIA-MetaFA-LSSVR proposto rendeu alta precisão e confiabilidade em previsões de 1 dia de antecedência do consumo de energia. Em particular, o sistema obteve valores médios de R, RMSE, MAE, MAPE, MaxAE e TER de 0,799, 0,161 kW h, 0,026 kW h, 14,785%, 0,185 kW h e 2,742%, respectivamente. Além disso, os resultados de comparação revelaram que o sistema de predição proposto é superior ao método de previsão linear (ou seja, SARIMA) e métodos de previsão não lineares (ou seja, LSSVR e MetaFA-LSSVR), comprovando o mérito de combinar modelos lineares e não lineares para análise de dados de consumo de energia em séries temporais. Notavelmente, o sistema apresentou melhores medidas de desempenho na faixa de 36,8 a 113,2% em comparação com os outros modelos.

Overall, o sistema proposto pode analisar eficientemente os dados coletados em tempo real de uma rede de rede inteligente. Através deste sistema, os usuários podem aumentar a eficiência energética de aparelhos e equipamentos elétricos em seus edifícios. O sistema pode ser potencialmente escamoso no gerenciamento de big data relacionados ao consumo de energia em rede inteligente. Em estudos futuros, um sistema dinâmico de otimização multiobjetiva será usado para otimizar o consumo de energia com base em padrões de uso para fornecer aos usuários horários alternativos de operação de aparelhos. Os usuários podem então comparar as alternativas para determinar quando os aparelhos devem ser ligados ou desligados. Um sistema de suporte a decisões de economia de energia baseado na Web será desenvolvido para fornecer aos usuários visualizações de insights úteis.

Neste estudo, os autores realizaram a análise de sensibilidade utilizando todos os atributos de dados coletados da infraestrutura de rede inteligente. A influência de fatores como a radiação solar não foi considerada na previsão do consumo de energia nesta investigação. Sugere-se, assim, um estudo mais aprofundado para coletar tais dados para possivelmente melhorar o desempenho preditivo.

Referências

- Yuan J, Hu Z. Desenvolvimento de eletricidade de baixo carbono na China – um perspective irsp baseado em rede super inteligente. *Renovar rev de energia sustentável* 2011;15:2707-13.
- Grid 2030 – vision nacional para o segundo 100 anos de eletricidade. Departamento de Energia dos Estados Unidos; 2003.
- Crisostomi E, Gallicchio C, Micheli A, Raugi M, Tucci M. Previsão das aplicações italianas de preço de eletricidade for smart grid. *Neurocomputing* 2015;170:286-95.
- Di Santo KG, Kanashiro E, Di Santo SG, Saidel MA. Uma revisão sobre redes inteligentes e experiências no Brasil. *Renovar o Sustain Energy Rev* 2015;52:1072-82.
- Soares F, Fotouhi Ghazvini MA, Vale Z, de Moura Oliveira PB. Um modelo multi-objetivo para o agendamento de recursos energéticos diários de uma rede inteligente com alta penetração de cargas sensíveis. *Appl Energy* 2016;162:1074-88.
- Colak I, Fulli G, Sagirolu S, Yesilbudak M, Covrig C-F. Projetos de smart grid na Europa: status atual, maturidade e cenários futuros. *Appl Energy* 2015;152:58-70.
- IEC. IEC Smart Grid Standardization Roadmap Edition 1.0. Grupo Estratégico da Grade Inteligente SMB; 2010.
- Kolokotsa D. O papel das redes inteligentes no setor de construção. *Construção de Energia* 2016;116:703-8.
- Cui B, Gao D-c, Wang S, Xue X. Análise de custo-benefício do ciclo de vida de armazenamentos a frio ativos para a construção de management demanda para aplicações de rede inteligente. *Appl Energy* 2015;147:523-35.
- Xue X, Wang S, Sun Y, Xiao F. Uma estratégia de gerenciamento de demanda de energia de construção interativa para facilitar a otimização da rede inteligente. *Appl Energy* 2014;116:297-310.
- Ahmad AS, Hassan MY, Abdullah MP, Rahman HA, Hussin F, Abdullah H, et al. Uma revisão sobre as aplicações de ANN e SVM para a construção de previsão de consumo de energia elétrica. *Renovar o Sustain Energy Rev* 2014;33:102-9.
- Zhang W, Liu S, Li N, Xie H, Li X. Desenvolvem análise de roteiro de previsão e tecnologia de energia renovável em edifícios na China. *Renovar o Sustain Energy Rev* 2015;49:395-402.
- Hsu D. Comparação de métodos integrados de agrupamento para previsão precisa e estável de dados de consumo de energia. *Appl Energy* 2015;160:153-63.
- Lü X, Lu T, Kibert CJ, Viljanen M. Modeling e previsão de consumo de energia para edifícios heterogêneos usando uma abordagem físico-estatística. *Appl Energy* 2015;144:261-75.
- Fumo N, Rafe Biswas MA. Análise de regressão para previsão de consumo de energia residencial. *Renovar o Sustain Energy Rev* 2015;47:332-43.
- Ekici BB, Aksoy UT. Previsão do consumo de energia por meio de networks neurais artificiais. *Adv Eng Softw* 2009;40:356-62.
- D'Andrea E, Lazzarini B, del Rosario SL. Previsão neural de consumo de energia devido à iluminação elétrica em edifícios de escritórios. *Sustentar a TIC da Internet Sustain* 2012;10:1-5.
- Li Q, Meng Q, Cai J, Yoshino H, Mochida A. Aplicando máquina vetorial de suporte para prever a carga de resfriamento por hora no edifício The. *Appl Energy* 2009;86:2249-56.
- Haydari Z, Kavehnia F, Askari M, Ganbariyan M. Modelagem de carga da série time e previsão de carga usando técnicas neuro-fuzzy. In: 9ª Conferência Internacional sobre qualidade e utilização de energia elétrica; 2007. p. 1-6.
- Naji S, Shamshirband S, Basser H, Keivani A, Alengaram UJ, Jumaat MZ, et al. Aplicação de metodologia adaptativa neuro-fuzzy para estimar o consumo de energia predial. *Renovar o Sustain Energy Rev* 2016;53:1520-8.
- Guo JJ, Wu JY, Wang RZ. Uma nova abordagem para a previsão de consumo de energia do aquecedor de água da bomba de calor doméstica com base na teoria do sistema cinza. *Construção de Energia* 2011;43:1273-9.
- Hamzacebi C, Es HA. Prever o consumo anual de eletricidade da Turquia usando um modelo cinza otimizado. *Energia* 2014;70:165-71.
- Box GEP, Jenkins GM. Análise da série time: previsão e controle. 3ª ed. Califórnia, Estados Unidos: Holden-day; 1970.
- Tan Z, Zhang J, Wang J, Xu J. Day-ahead preço de eletricidade prevendo usando wavelet transform combinado com models ARIMA e GARCH. *Appl Energy* 2010;87:3606-10.
- Caixa GEP, Jenkins G. Análise da série time, previsão e controle. Califórnia, Estados Unidos: Holden-Day, Incorporated; 1990.
- Khashei M, Bijari M. Uma nova hibridização de redes neurais artificiais e modelos ARIMA para previsão de séries temporais. *Appl Soft Comput* 2011;11:2664-75.
- Wang Y, Wang J, Zhao G, Dong Y. Aplicação de abordagem de modificação residual em ARIMA sazonal para previsão de demanda de eletricidade: um estudo de caso da China. *Política Energética* 2012;48:284-94.
- Pal M, Deswal S. Support vector regression based shear strength modelagem de vigas profundas. *Comput Struct* 2011;89:1430-9.
- Martins FF, Begonha A, Amália Sequeira Braga M. Previsão do comportamento mecânico do granito do Porto utilizando técnicas de Mineração de Dados. *Expert Syst Appl* 2012;39:8778-83.
- De Boeck L, Verbeke S, Audenaert A, De Mesmaeker L. Melhorando o desempenho energético dos edifícios residenciais: uma revisão da literatura. *Renovar o Sustain Energy Rev* 2015;52:960-75.
- Fouquier A, Robert S, Suard F, Stéphan L, Jay A. State of the art in building modelling and energy performances forecast: a review. *Renovar o Sustain Energy Rev* 2013;23:272-88.
- Fumo N. Uma revisão sobre o básico da estimativa de energia de construção. *Renovar o Sustain Energy Rev* 2014;31:53-60.
- Li K, Hu C, Liu G, Xue W. Building's electricity consumption forecast using optimized artificial neural networks and principal component analysis. *Construção de Energia* 2015;108:106-13.
- Mena R, Rodríguez F, Castilla M, Arahal MR. Um modelo de prediction baseado em redes neurais para o consumo de energia de um edifício bioclimático. *Construção de Energia* 2014;82:142-55.
- Chen H, Huang Y, Shen H, Chen Y, Ru M, Chen Y, et al. Modelagem de variações temporais no consumo global de energia residencial e emissões de poluentes. *Appl Energy* 2015. <http://dx.doi.org/10.1016/j.apenergy.2015.10.185>.
- Fan C, Xiao F, Wang S. Desenvolvimento de modelos de previsão para o consumo de energia de construção do próximo dia e pico de demanda de energia usando técnicas de mineração de dados. *Appl Energy* 2014;127:1-10.
- Zhao H-X, Magoulès F. Uma revisão sobre a previsão de construção do consumo de energia. *Renovar a Energia Sustentável Rev* 2012;16:3586-92.
- Sandels C, Widén J, Nordström L, Andersson E. Previsões antecipadas de consumption de eletricidade em um edifício de escritórios suecos a partir de dados meteorológicos, ocupações e temporais. *Construção de Energia* 2015;108:279-90.
- Zhang PG. Previsão da série temporal usando um modelo híbrido de ARIMA e rede neural. *Neurocomunicação* 2003;50:159-75.
- Hyndman RJ, Athanasopoulos G. Previsão: princípios e prática; <<https://www.otexts.org/fpp/Otexts>>.
- Chen K-Y, Wang C-H. Um SARIMA híbrido e máquinas vetoriais de suporte na previsão dos valores de produção da indústria de máquinas em Taiwan. *Expert Syst Appl* 2007;32:254-64.
- Choi T-M, Yu Y, Au K-F. Um método híbrido de transformação de ondas SARIMA para previsão de vendas. *Decis Support Syst* 2011;51:130-40.
- Jung HC, Kim JS, Heo H. Previsão do consumo de energia de construção usando um algoritmo genético codificado real melhorado baseado em pelo menos quadrados de suporte à abordagem da máquina vetorial. *Construção de Energia* 2015;90:76-84.
- Lee Y-S, Tong L-L. Previsão de séries temporais não lineares de consumo de energia usando um modelo dinâmico híbrido. *Appl Energy* 2012;94:251-6.
- Sánchez Lasheras F, de Cos Juez FJ, Suárez Sánchez A, Krzemien' A, Riesgo Fernández P. Previsão do preço spot de cobre COMEX por meio de redes neurais e modelos ARIMA. *Resour Policy* 2015;45:37-43.
- Khandelwal I, Adhikari R, Verma G. Time series forecasting using hybrid ARIMA and ANN models based on DWT decomposition. *Procedia Comput Sci* 2015;48:173-9.
- Gp de Zhang. Previsão de séries temporais usando um modelo híbrido de ARIMA e rede neural. *Neurocomunicação* 2003;50:159-75.
- Hansen JV, Nelson RD. Análise da série time com networks neurais e híbridos de rede ARIMAneural. *J Exp Theor Artif Intell* 2003;15:315-30.

- [49] Sallehuddin R, Hj. Shamsuddin SM. Rede neural artificial cinzenta híbrida e modelo médio móvel integrado auto regressivo for previsão de dados de séries temporais. *Appl Artif Intell* 2009;23:443-86.
- [50] Pai P-F, Lin C-S. Um modelo híbrido de ARIMA e máquinas vetoriais de suporte na previsão de preços das ações. *Ômega* 2005;33:497-505.
- [51] Da-yong Z, Hong-wei S, Pu C. Stock modelo de previsão de mercado baseado em um ARMA híbrido e máquinas vetoriais de suporte. In: 15º Processo anual de conferência, internationalconferenceonmanagementscience andengineering;2008.p.1312-7.
- [52] Che J, Wang J. Preços de eletricidade de curto prazo previstos com base na regressão vetorial de suporte e modelagem média móvel integrada auto-regressive. *Energy Convers Manage* 2010;51:1911-7.
- [53] Su S, Zhang W, Zhao S. Previsão de falha para sistema não linear usando ARMA deslizante combinado com LS-SVR on-line. *Math Probl Eng* 2014;2014:9.
- [54] Haifeng W, Dejin H. Comparação de SVM e LS-SVM para regressão. In: Conferência internacional sobre redes neurais e cérebro; 2005. p. 279-83.
- [55] Suykens JAK, Gestel TV, Brabanter JD, Moor BD, Vandewalle J. Least quadrados suportam máquinas vetoriais. Cingapura: Ciência Mundial; 2002.
- [56] Chou J-S, Ngo N-T, Pham A-D. Previsão de força de corte em vigas profundas de concreto armado usando regressão metaéstric de suporte metahúscrista inspirada na natureza. *J Comput Civil Eng* 2016;30:04015002.
- [57] Yang X-S. Capitulo 2 – Análise de algoritmos. Algoritmos de otimização inspirados na natureza. Oxford, Reino Unido: Elsevier; 2014. p. 23-44.
- [58] Roque CMC, Martins PALS. Evolução diferencial para otimização de feixes de grau funcional. *Compos Struct* 2015;133:1191-7.
- [59] Fister I, Fister Jr I, Yang X-S, Brest J. Uma revisão abrangente dos algoritmos do vagalume. *Swarm Evolut Comput* 2013;13:34-46.
- [60] Talatahari S, Gandomi AH, Yun GJ. Design ideal de estruturas de torres usando algoritmo firefly. *Struct Des Tall Spec Build* 2014;23:350-61.
- [61] Tseng F-M, Tzeng G-H. Um modelo ARIMA sazonal difuso para previsão. *Fuzzy Sets Syst* 2002;126:367-76.
- [62] Alwee R, Hj Shamsuddin SM, Sallehuddin R. Hybrid suportam regressão vetorial e modelos de média móvel integrada autoregressive melhorados pela otimização de exame de partículas para ratos de crime patrimonial previstos com indicadores econômicos. *Sci World J* 2013;2013:11.
- [63] Shamshirband S, Mohammadi K, Yee PL, Petkovic' D, Mostafaeipour A. Uma avaliação comparativa para identificar a adequação da máquina de aprendizagem extrema para prever radiação global global de solar horizontal. *Renovar o Sustain Energy Rev* 2015;52:1031-42.
- [64] Raza MQ, Khosravi A. Uma revisão sobre técnicas de previsão de demanda de carga baseada em inteligência artificial para redes inteligentes e edifícios. *Renove o Sustain Energy Rev* 2015;50:1352-72.
- [65] Vapnik VN. A natureza da teoria da aprendizagem estatística. Nova Iorque, Estados Unidos: Springer-Verlag, Nova Iorque; 1995.
- [66] Yang X-S. Algoritmo firefly. Bristol, Reino Unido: Luniver Press; 2008.
- [67] Coelho LDS, Mariani VC. Abordagem aprimorada do algoritmo de vagalume aplicado ao carregamento de refrigerador para conservação de energia. *Construção de Energia* 2013;59:273-8.
- [68] He D, He C, Jiang LG, Zhu HW, Hu GR. Características caóticas de um mapa iterativo unidimensional com colapsos infinitos. *IEEE Transact Circ Syst I: Fundam Theory Appl* 2001;48:900-6.
- [69] Chou J-S, Ngo N-T. Estrutura de análise de dados de grade inteligente para aumentar a economia de energia em edifícios residenciais. *Autom Constr* 2016. <http://dx.doi.org/10.1016/j.autcon.2016.01.002>.
- [70] Chou J-S, Telaga AS. Detecção em tempo real do consumo de energia anômeos. *Renovar o Sustain Energy Rev* 2014;33:400-11.