

ANALYSIS AND COMPARISON BETWEEN REGRESSION MODELS FOR TEMPERATURE ESTIMATION OF SOLAR COLLECTORS OPERATING WITH NANOFUIDS

Juarez Pompeu de Amorim Neto
Paulo Alexandre Costa Rocha
Felipe Pinto Marinho
Ricardo José Pontes Lima
Lino Wagner Castelo Portela
Maria Eugênia Vieira da Silva
juareznet33@hotmail.com
paulo.rocha@ufc.com
felipe.pinto.marinho@gmail.com
rjponteslima@gmail.com
eugenia@les.ufc.br

Mechanical Engineering Department, Technology Center, Pici, Federal University of Ceara, Fortaleza - CE, Zip Code 60440-593, Brazil.

Abstract. The objective of this work is to verify the application of polynomial regression methods, Ridge and Lasso regression in the nowcasting of the fluid temperature and energy gain of a solar collector operating with nanofluids. The collector has temperature and global/direct solar radiation sensors for data logging. In addition the R programming language was used for the statistical analysis of R^2 , MAE (Mean Absolute Error) and RMSE (Root Mean Squared Error). The models were applied in three different data sets, which regarded to the data for water temperature prediction and TiO₂ nanofluids with a concentration of 25 ppm and 75 ppm, where each method applied seven predictors for the fluid temperature nowcasting. The best Root Mean Squared error found in the test sets was 2.281°C for a degree 3 polynomial regression, whereas the Ridge presented an RMSE of 3.190°C. The Ridge and the Lasso usually improve least squares methods but they did not perform well in this data set, the Ridge regression considered a model with all the predictors and got a high test error, as far as the Lasso excluded some predictors and got an improved result. A cross-validation was performed to know the degree of the most effective polynomial for the analysis of these data and the polynomial regression of degree 3 obtained the best result, confirming that the fluid temperature does not follow a linear trend mainly during the hours from 5:30 to 21:30.

Keywords: Solar Energy, Renewable Energy, Machine Learning, Ridge Regression, LASSO.

1 Introdução

Ultimamente, o uso das energias renováveis tem aumentado devido ao excessivo consumo de combustíveis fósseis que contribuem para a poluição ambiental. Isso vem se tornando um fator preocupante principalmente por conta da industrialização e globalização das sociedades modernas, Farajzadeh, Movahed e Hosseini [1]. Desde 1990, as fontes de energias renováveis têm a maior taxa de crescimento em comparação com outras fontes, em especial a energia solar, Mwesigye et al. [2], nesse contexto, o uso da energia solar térmica, que é uma fonte barata, vem sendo bastante utilizada para aproveitar a energia solar usando em aquecimentos de água para uso doméstico, aplicações industriais e comerciais. Os coletores solares térmicos são usados para captar essa fonte de energia, onde a sua superfície absorve a energia solar útil e depois transfere o calor para o fluido de trabalho, sendo assim a eficiência desses coletores dependem de como o absorvedor capta a energia e a forma como o calor é transferido para o fluido, onde geralmente têm baixa eficiência de conversão energética, Bandarra Filho et al. [3].

Visando melhorar a eficiência desses coletores, uma solução viável seria mudar o fluido de trabalho, geralmente a água por nanofluidos, Ebrahimi-moghadam et al. [4]. Os nanofluidos são nanopartículas geralmente de (1 – 100 nm) que são adicionadas no fluido de trabalho e melhoram a capacidade de absorção, Devendiran e Amirtham [5]. Diversos tipos de nanofluidos vem sendo usados em pesquisas recentes, tais como a prata em Bandarra Filho et al. [3], Al_2O_3 em Alade et al. [6], CuO em Michael e Iniyar [7]. Alguns trabalhos mostraram como os nanofluidos são eficazes quando adicionados no fluido base, melhorando de forma significativa a eficiência de absorção energética, como por exemplo o nanofluido de ouro que melhorou em 76,0% a eficiência de conversão energética em Jin et al. [8], o nanofluido SWCNTs-Therminol@VP-1 melhorou em 234% em Mwesigye, Yilmaz e Meyer [2], a prata que aprimorou em 144% em Bandarra Filho et al. [3] e CeO_2 que aumentou em 42,3% em Sharafeldin e Gróf [9]. Os trabalhos utilizando nanopartículas no fluido base mostraram que o aumento da concentração implica em uma maior transferência de calor para o nanofluido.

Visando obter o melhor aproveitamento do fluido de trabalho, é importante identificar os parâmetros físicos que podem influenciar no desempenho do coletor solar, tais como a radiação solar global, radiação solar direta, radiação solar difusa, temperatura ambiente, sensação térmica, hora e a data. Diversos métodos numéricos são capazes de identificar a relação existente entre a capacidade de conversão energética e os parâmetros em análise. Uma das principais vantagens da simulação numérica é o custo, em contrapartida os resultados podem não ser satisfatórios quando comparado ao desempenho real, Liu et al. [10], em função disso, é importante a análise e comparação de diversos métodos que possam prever o comportamento do fluido em um coletor. O aprendizado de máquina é um dos campos da ciência da computação, podendo ser aplicado em diversas áreas, encontrando relações entre variáveis de entradas e saídas, tornando possível o reconhecimento de padrões e estabelecendo previsões, Voyant et al. [11].

Neste trabalho, os métodos Ridge, Lasso e a regressão polinomial foram usados para prever a temperatura do fluido no coletor solar e analisar as formas de absorção de calor dos mesmos. Verificou-se também influência dos parâmetros físicos na temperatura do fluido de trabalho em um coletor localizado na cidade de Fortaleza – CE.

2 Aprendizagem de máquina

Esse tópico detalha o funcionamento de cada algoritmo usado para predição da temperatura do fluido de trabalho.

2.1 Ridge

A regressão Ridge busca normalizar os preditores, onde os atributos ficam na mesma escala com

desvio padrão igual a 1, além de minimizar a seguinte função de custo descrita pela Eq. (1) na qual a variável y representa a resposta, x são os preditores e β representam os coeficientes de regressão.

$$\sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^p \beta_j^2. \quad (1)$$

Onde $\|\beta\|_2$ é a norma L2 da segunda parcela da equação.

A primeira parcela é usada na regressão linear e é chamada de RSS (soma dos resíduos quadrados) como em Fang e Lahdelma [12]. O parâmetro λ , também chamado de penalidade de contração, deve ser determinado para que a equação (1) obtenha o menor valor possível, caso λ seja igual a zero, o problema se resume a um problema de mínimos quadrados. Para cada valor de penalidade de contração há uma estimativa para os coeficientes de regressão, ao contrário da regressão por mínimos quadrados que só gera apenas um conjunto de estimativas. A normalização dos preditores é dada pela Eq. (2):

$$\tilde{x}_{ij} = \frac{x_{ij}}{\sqrt{\frac{1}{n} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}}. \quad (2)$$

Onde o denominador é o desvio padrão estimado do preditor j .

À medida que o termo de penalidade aumenta, os valores dos coeficientes diminuem, levando a diminuição da variância e aumento da tendência, já para um valor baixo de custo, tem-se um aumento na variância e redução da tendência. A vantagem da regressão Ridge em comparação com os mínimos quadrados está na relação entre tendência e variância, (*trade-off* tendência - variância), já que o erro médio quadrático é função da variância mais a tendência ao quadrado, conforme JAMES et al. [13]. Para a escolha do melhor valor de λ , é feita uma validação cruzada que obtenha o menor erro de validação.

2.2 Lasso

A regressão de Ridge tem uma desvantagem bem explícita, onde ao minimizar a função objetivo o custo λ apenas irá reduzir os valores dos coeficientes, nunca se igualando a zero. A regressão Lasso busca solucionar esse problema, podendo reduzir os coeficientes de regressão para zero, sendo um algoritmo que selecionará os preditores que mais influenciam na resposta, Yang e Wen [14]. A regressão Lasso se dá por meio da minimização da seguinte função de custo representada pela Eq. (3).

$$RSS + \lambda \sum_{j=1}^p |\beta_j|. \quad (3)$$

Onde $\|\beta\|_1 = \sum_{j=1}^p |\beta_j|$ é norma L₁, também usada em He et al. [15].

O processo de encolhimento dos coeficientes além de selecionar os preditores, permite uma melhor interpretação do modelo analisado. De maneira análoga, a regressão Lasso produzirá diversas soluções para os coeficientes de regressão, onde será selecionado o conjunto que obtiver o menor erro de validação ao selecionar o parâmetro λ .

2.3 Regressão Polinomial

Em busca de um método que conseguisse resultados satisfatórios em um conjunto de dados cuja relação entre os preditores e a resposta fosse não linear, buscou-se substituir o modelo linear padrão pelo modelo representado pela Eq. (4).

$$y = \varepsilon + \beta_0 + \sum_{i=1}^n \beta_i x^i. \quad (4)$$

Onde ε é o termo associado ao erro da regressão polinomial.

O grau da regressão polinomial é dado pelo maior expoente aplicado no preditor, onde com um maior valor de n em Eq. (1), tem-se uma curva mais flexível e com um formato incomum, sendo geralmente usado um grau de polinômio 3 ou 4 para garantir resultados satisfatórios como afirma JAMES et al. [13]. Uma validação cruzada é feita para escolher o grau do polinômio que melhor se ajusta ao conjunto de dados.

3 Descrição dos dados

Nessa fase, é detalhada a divisão dos conjuntos de treino e teste, além de mostrar como foi feita a validação cruzada do conjunto de dados. A sessão também aborda como foram coletados os dados e quais preditores foram usados para avaliação dos métodos.

3.1 Organização do conjunto de dados

Neste trabalho, o conjunto de dados foi dividido em duas partes, uma parte denominada conjunto de treinamento e a outra parte de conjunto de teste. O conjunto de treino é composto de 70% das observações dispostas de forma aleatória e o restante faz parte do conjunto de teste como em Ahmad, Reynolds e Rezgui [16]. Os dados de treinamento foram utilizados para treinar os métodos calculando os hiperparâmetros ideais para o modelo, onde a escolha desses parâmetros livres é um processo bem criterioso. Visando obter a solução ideal, uma validação cruzada 10-folds foi realizada no conjunto de treinamento, onde as observações foram divididas em 9 subconjuntos e a validação cruzada foi realizada no subconjunto restante, esse processo foi repetido para cada subconjunto (10 vezes). Esse processo foi realizado para que os modelos conseguissem bom desempenho em dados desconhecidos. Logo após a etapa de treinamento com a escolha dos melhores hiperparâmetros, o conjunto de teste foi utilizado para validar a modelagem matemática. A Fig. 1 mostra o processo de treinamento e de teste dos dados.

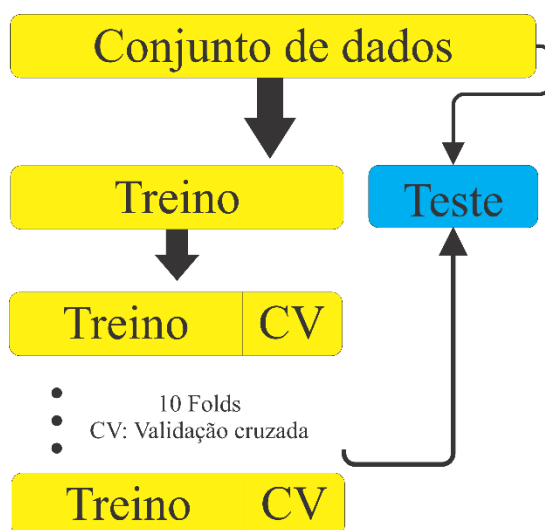


Figura 1. Divisão e validação do conjunto de dados

Para uma análise de desempenho dos modelos, algumas métricas foram analisadas para verificar se o modelo se comportava de forma satisfatória no conjunto de teste. A estatística MAE (Erro Absoluto Médio): magnitude média dos erros de previsão, o RMSE (Raiz do Erro Médio Quadrático): erro médio de previsão que atribui um peso aos erros mais altos e o R^2 foram as métricas de análise dos modelos, onde os cálculos são mostrados na Eq. 5 e Eq. 6. Essas métricas foram usadas em trabalho similares para avaliar a precisão de outros modelos, Alfadda et al. [17], Wang et al. [18], Chu e Coimbra [19], Nonnenmacher et al. [20].

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}} \quad (5)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i| \quad (6)$$

3.2 Coleta de dados

O sistema solar térmico que está localizado na cidade de Fortaleza (Latitude: 03° 43' 02" Sul, Longitude: 38° 32' 35" Oeste) foi utilizado para captação dos dados de temperatura da água no período de 11/08/2018 – 21/09/2018, o mesmo possui uma estrutura de alumínio com placas de vidros com espessura de 4 mm, comprimento de 1000 mm e altura de 500 mm.

Para aquisição dos dados de temperatura foram utilizados 6 sensores do tipo K que medem a temperatura da água, um mesmo sensor que mede a temperatura ambiente e outro sensor que mede a sensação térmica. Foram coletados também os dados de radiação solar global e difusa por meio de um piranômetro e os dados de radiação solar direta por meio de um pireliômetro. Os dados de entrada da modelagem numérica são a temperatura ambiente, sensação térmica, radiação solar direta, radiação solar difusa e radiação solar global, todos eles foram usados para prever a temperatura do fluido de trabalho. A Fig. 2 mostra a estrutura e os sensores utilizados no trabalho.

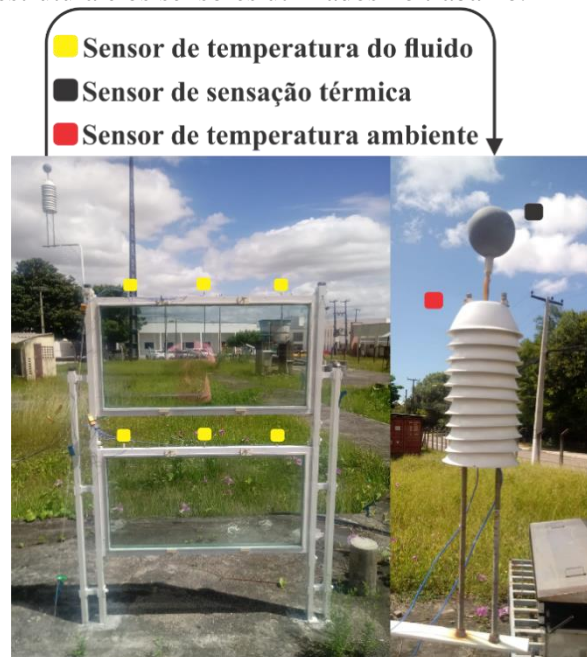


Figura 2. Sistema para coleta de dados

4 Resultados e discussões

Essa sessão irá detalhar os resultados dos três tipos de algoritmos de aprendizagem de máquina, onde a análise será focada no desempenho do modelo no conjunto de teste em função das métricas RMSE, MAE e R^2 .

4.1 Ridge

A regressão Ridge produz diversos conjuntos de coeficientes em função do parâmetro livre λ que minimiza a função de custo. Durante a validação cruzada no conjunto de dados de treinamento, variou-se λ para encontrar aquele que obtivesse o menor erro de validação cruzada. A fig. 3 mostra o processo de treinamento para os dados de água que busca o determinado hiperparâmetro, onde com um valor para λ de 0,5615 obteve-se o menor RMSE de 3,0310, concluindo assim a etapa de busca do valor de λ para a água. A tabela 1 expõe os resultados de validação cruzada para o nanofluido a 25 ppm e 75 ppm.

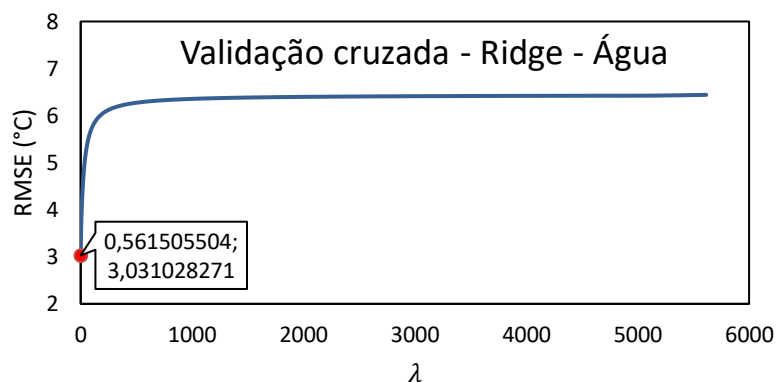


Figura 3. Busca do Hiperparâmetro

Tabela 1. Validação cruzada Ridge

TiO ₂ – 25 ppm		TiO ₂ – 75 ppm	
$\lambda_{\text{ótimo}}$	RMSE (°C)	$\lambda_{\text{ótimo}}$	RMSE (°C)
0,6448	2,6624	0,6059	2,9640

Após a etapa de treinamento o modelo ajustado selecionado foi utilizado para tentar prever o restante do conjunto de dados (etapa de teste) e a fig. 4 mostra o resultado que compara os valores previstos e os valores mensurados, onde percebe-se um resultado muito inconsistente, já que os valores estimados não conseguem seguir com tanta precisão os valores medidos, o método de predição oscila bastante no decorrer das observações obtendo uma dispersão considerável. Conseguem-se observar também que em um pico e dois vales, o algoritmo se distancia muito do valor real. O valor baixo do parâmetro λ causa um aumento na variância, explicando o fato das previsões oscilarem muito.

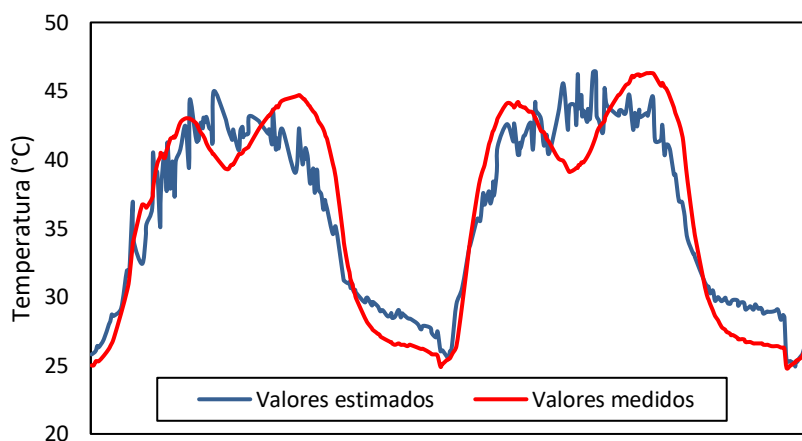


Figura 4. Predição - Ridge

A tabela 2 mostra os resultados finais do aprendizado, expondo os erros de teste e treinamento. Infere-se do resultado que a modelagem Ridge produz métricas com valores não satisfatórios, onde um menor valor de RMSE na fase de teste de 2,6527 °C representa a diferença entre o valor real e o previsto para o nanofluido a 25 ppm e esse ainda é um valor alto para essa diferença. Al-Obeidat et al. [21] utilizou a regressão Ridge para previsão de temperaturas em intervalos futuros e encontrou um RMSE para o mesmo método de 2,1211°C, conseguindo reduzir o erro em 20,03% em comparação com esse trabalho.

Tabela 2. Erros de teste e treinamento – Ridge

Conjunto de dados - Ridge	Treinamento			Teste		
	RMSE	MAE	R^2	RMSE	MAE	R^2
Ridge - água	3,0236	2,5855	0,7795	6,0477	2,7093	0,1189
Ridge – 25 ppm	2,5909	2,1342	0,8700	2,6527	2,1715	0,8591
Ridge – 75 ppm	2,9517	2,4610	0,8140	3,1191	2,6206	0,8088

4.2 Lasso

De maneira análoga a regressão Ridge, a regressão Lasso também tem o objetivo de determinar o melhor valor do hiperparâmetro que minimiza a função de custo. Os resultados de validação cruzada para a água é mostrado na fig. 5, onde com $\lambda = 0,0069$ chegou-se em um RMSE de 2,8630. Na tabela 3 são expostos os mesmos resultados para o conjunto de dados de nanofluido a 25 ppm e 75 ppm.

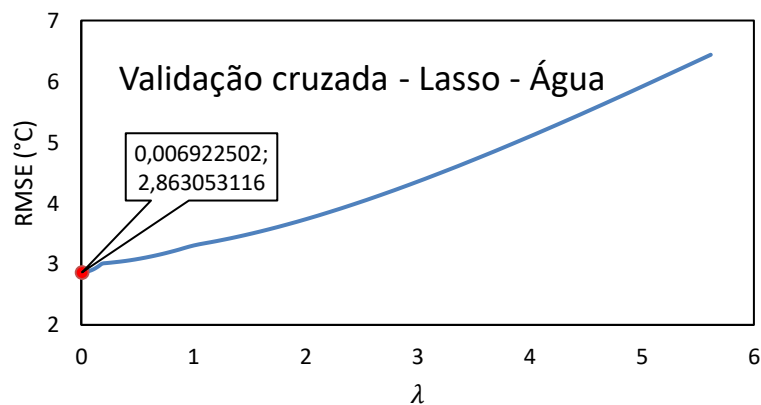


Figura 5. Busca do Hiperparâmetro

Tabela 3. Validação cruzada – Lasso

TiO ₂ – 25 ppm		TiO ₂ – 75 ppm	
$\lambda_{\text{ótimo}}$	RMSE (°C)	$\lambda_{\text{ótimo}}$	RMSE (°C)
0,0105	2,5549	0,0131	2,8850

A etapa de teste foi realizada aplicando o modelo ajustado ao restante dos dados, idêntico ao método Ridge, onde o resultado pode ser visto na fig. 6 que apresenta resultados bem similares ao Ridge, já que ambos resultaram em um valor de λ pequeno apresentando alta variância e não conseguindo obter uma previsão equivalente com os valores medidos.

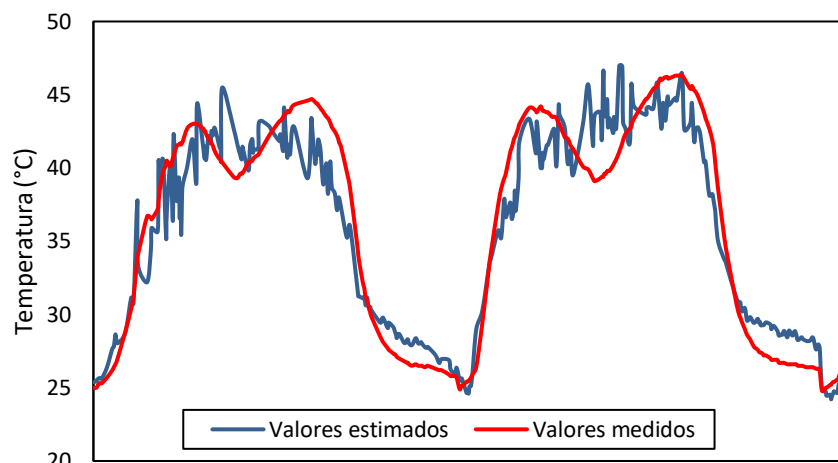


Figura 6. Predição – Lasso

Lasso também funciona como um seletor de preditores, excluindo aqueles que menos influenciam no resultado, a tabela 4 mostra os coeficientes de regressão e os recursos utilizados. O resultado mostra que nas três situações a radiação solar difusa é excluída, sendo um preditor irrelevante para o modelo. Nos dados de água além da radiação solar difusa, Lasso excluiu a data. Outro fator relevante na seleção dos preditores é que a temperatura ambiente atinge os valores mais altos nas 3 situações, sendo o preditor que mais interfere na resposta do modelo preditivo.

Tabela 4. Seleção de preditores

Coeficientes	Água	TiO ₂ – 25 ppm	TiO ₂ – 75 ppm
Termo independente	-51,6890	-52,2374	-44,0021
Data	-	0,0288	0,0084
Hora	-0,0738	-0,3480	0,0619
Temperatura ambiente	2,2199	3,2583	1,6478
Sensação térmica	-0,7641	0,0156	0,9956
Radiação difusa	-	-	-
Radiação Global	-0,0052	-0,0013	-0,0032
Radiação direta	-0,0004	-0,0040	-0,0017

A tabela 5 mostra o resultado final para o conjunto de treinamento e teste em função das métricas apresentadas. Os resultados mostram que o Lasso conseguiu uma pequena melhoria nos resultados de teste em comparação com o Ridge, sendo 4,15% melhor em função da análise da métrica RMSE. Percebe-se então que selecionar os preditores pode causar uma leve melhoria no resultado quando se compara com a regressão Ridge.

Tabela 5. Erros de teste e treinamento – Lasso

Conjunto de dados - Lasso	Treinamento			Teste		
	RMSE	MAE	R ²	RMSE	MAE	R ²
Lasso - água	2,8590	2,3688	0,8029	4,0715	2,4357	0,6006
Lasso – 25 ppm	2,4082	1,9200	0,8877	2,5425	1,9962	0,8706
Lasso – 75 ppm	2,8660	2,3656	0,8246	2,9339	2,4328	2,8308

Em ambos algoritmos o conjunto de dados de nanofluido 25 ppm obteve o menor RMSE de teste, seguido pelo nanofluido 75 ppm e com os piores valores de RMSE de teste, a água. Al-Obeidat et al. [21] obteve um RMSE de 2,0898 °C, obtendo uma redução de 17,80% em relação a esse trabalho.

4.3 Regressão polinomial

A regressão polinomial busca o grau de polinômio que oferece o menor erro na fase de teste, uma validação cruzada foi feita no conjunto de treinamento em busca do grau polinomial que oferecesse o menor valor de RMSE e maior valor de R^2 , a fig. 7 mostra o resultado para o conjunto de dados da água, onde um menor valor de RMSE foi de 2,1719 e um maior valor de R^2 foi de 0,8868 para um grau polinomial 18, mas um alto grau de polinômio resultaria em uma função muito flexível, resultando em um ajuste excessivo nos dados de treinamento e podendo comprometer seu desempenho em dados desconhecidos (conjunto de teste). Visando solucionar tal problema, optou-se por errar mais na fase de treinamento para obter um melhor desempenho na fase de teste, escolhendo um grau de polinômio mais próximo da interseção das curvas, obtendo um valor de RMSE de 2,5474 e de R^2 de 0,8442 para um polinômio de grau 3. A tabela 6 mostra a validação cruzada para o conjunto de dados de nanofluido 25 ppm e 75 ppm.

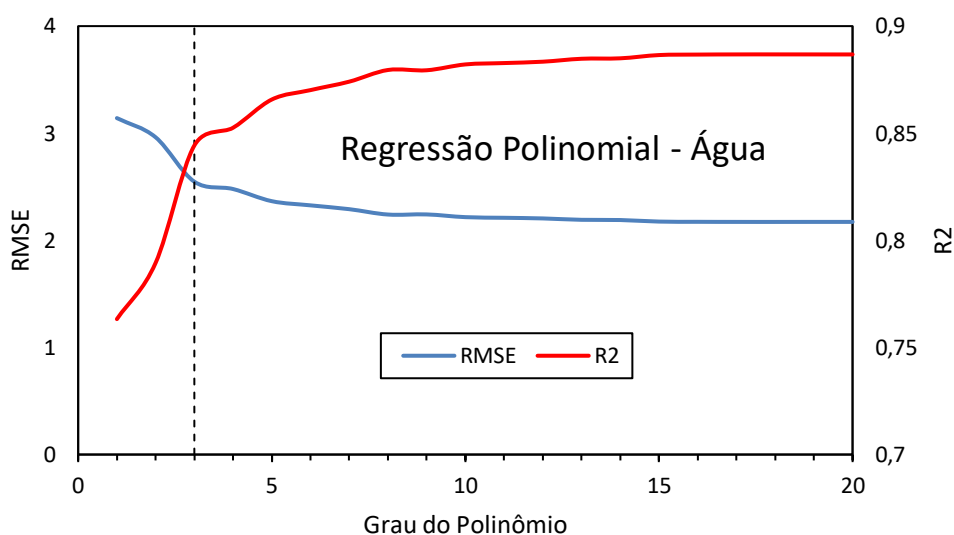


Figura 7. Escolha do grau do polinômio

Tabela 6. Validação cruzada – Regressão polinomial

TiO ₂ – 25 ppm			TiO ₂ – 75 ppm		
Grau do Polinômio	RMSE (°C)	R ²	Grau do Polinômio	RMSE (°C)	R ²
3	1,9347	0,9292	2	2,7021	0,8872

Após a fase de treinamento, o modelo ajustado foi testado nos 30% restante dos dados para verificar a acuracidade do método, a fig. 8 mostra o resultado. A regressão polinomial em algumas partes consegue acompanhar os dados medidos, mas em determinadas fases, onde há pico ou vale dos dados medidos, a regressão tem uma tendência constante com pequenas oscilações.

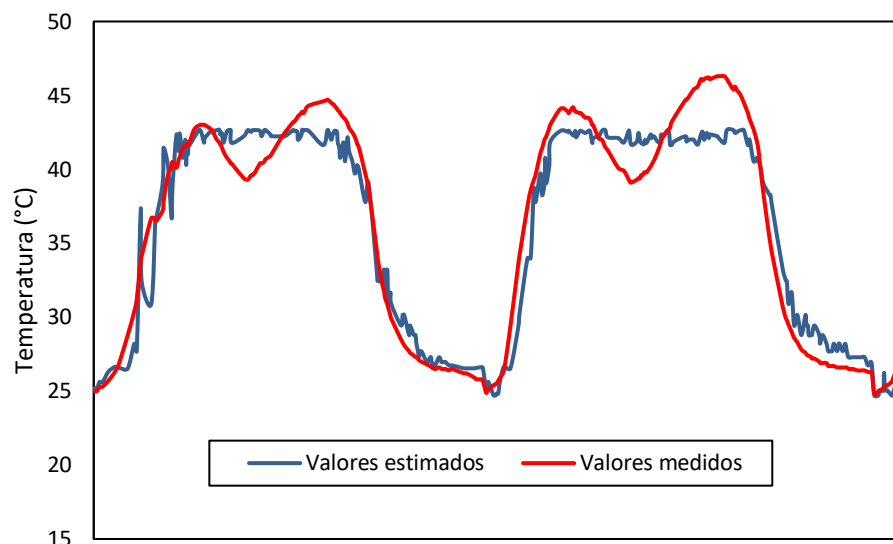


Figura 8. Predição – Regressão polinomial

Para quantificar essa predição foram calculadas as métricas de erros para os conjuntos de treino e teste, conforme a tabela 7. Percebe-se que a regressão polinomial consegue reduzir ainda mais os valores de RMSE nos conjuntos de teste em relação aos métodos de regressão linear, alcançando um valor de 1,7894 para o nanofluido 25ppm. O pior valor de RMSE (2,1837) para a regressão polinomial ainda consegue obter um desempenho superior aos métodos Ridge e Lasso, a fig. 9 mostra um comparativo entre os valores medidos e estimados para todos os métodos.

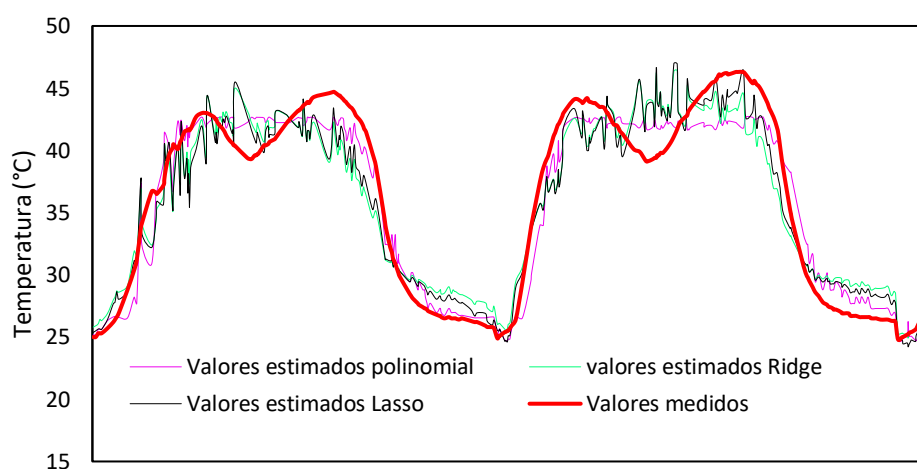


Figura 9. Comparação entre os métodos

Tabela 7. Erros de teste e treinamento – Regressão polinomial

Conjunto de dados de regressão	Treinamento			Teste		
	RMSE	MAE	R^2	RMSE	MAE	R^2
Água	2,1719	1,7381	0,8868	2,1837	1,7076	0,8851
25 ppm	1,7240	1,4004	0,9412	1,7894	1,4314	0,9380
75 ppm	1,8526	1,4178	0,9325	1,9729	1,4933	0,9174

5 Conclusões

O trabalho detalha o uso de algoritmos de aprendizagem de máquina para predição de temperatura de um sistema solar térmico localizado na cidade de Fortaleza. Os resultados foram capazes de simular o comportamento do coletor solar e até verificar a capacidade de absorção de nanofluidos, podendo até competir com modelos físicos, como os de balanço de energia.

Ridge e LASSO são os métodos que melhoram a regressão linear e foram avaliados nesse trabalho de predição de temperatura. A regressão Ridge que leva em conta todos os recursos de entrada foi o método com os resultados menos satisfatórios, obtendo um RMSE de teste de 2,6527 °C, já a regressão LASSO que seleciona os preditores mais importantes obteve um RMSE de teste de 2,5425 °C, onde essa regressão excluiu a radiação solar difusa no modelo final, infere-se então que a radiação solar difusa não se torna necessário para prever a temperatura da água no coletor solar, já que LASSO obteve uma pequena melhoria em comparação com Ridge.

A regressão polinomial foi outro método de avaliação nesse trabalho, o resultado do RMSE de teste (1,7894 °C) obteve uma melhoria significativa em relação aos métodos lineares, obtendo uma melhoria de 32,5% em relação ao Ridge e de 29,6% em relação ao Lasso. Isso mostra que a temperatura do fluido de trabalho do coletor não segue uma tendência linear, mas sim uma variação que pode ser prevista através de um polinômio de grau 3.

Os conjuntos de dados analisados nesse trabalho foram a água, nanofluido TiO₂ com 25 ppm e 75 ppm. O conjunto com a melhor performance nos três diferentes métodos foi o nanofluido TiO₂ a 25 ppm seguido de TiO₂ a 75 ppm e água respectivamente. A regressão polinomial se mostrou eficiente para predição da temperatura do fluido, esse resultado pode ser usado quando se deseje uma estimativa imediata da temperatura de trabalho do fluido.

AGRADECIMENTOS

Os autores agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

REFERÊNCIAS

- [1] FARAJZADEH, Ehsan; MOVAHED, Saeid; HOSSEINI, Reza. Experimental and numerical investigations on the effect of Al₂O₃/TiO₂H₂O nanofluids on thermal efficiency of the flat plate solar collector. *Renewable Energy*, [s.l.], v. 118, p.122-130, abr. 2018. Elsevier BV. <http://dx.doi.org/10.1016/j.renene.2017.10.102>.
- [2] MWESIGYE, Aggrey; YILMAZ, İbrahim Halil; MEYER, Josua P. Numerical analysis of the thermal and thermodynamic performance of a parabolic trough solar collector using SWCNTs-Therminol®VP-1 nanofluid. *Renewable Energy*, [s.l.], v. 119, p.844-862, abr. 2018. Elsevier BV. <http://dx.doi.org/10.1016/j.renene.2017.10.047>.
- [3] BANDARRA FILHO, Enio Pedone et al. Experimental investigation of a silver nanoparticle-based direct absorption solar thermal system. *Energy Conversion And Management*, [s.l.], v. 84, p.261-267, ago. 2014. Elsevier BV. <http://dx.doi.org/10.1016/j.enconman.2014.04.009>.
- [4] EBRAHIMI-MOGHADAM, Amir; MOHSENI-GHARYEHSAFA, Behnam; FARZANEH-GORD, Mahmood. Using artificial neural network and quadratic algorithm for minimizing entropy generation of Al₂O₃-EG/W nanofluid flow inside parabolic trough solar collector. *Renewable Energy*, [s.l.], v. 129, p.473-485, dez. 2018. Elsevier BV. <http://dx.doi.org/10.1016/j.renene.2018.06.023>.
- [5] DEVENDIRAN, Dhinesh Kumar; AMIRTHAM, Valan Arasu. A review on preparation, characterization, properties and applications of nanofluids. *Renewable And Sustainable Energy Reviews*, [s.l.], v. 60, p.21-40, jul. 2016. Elsevier BV. <http://dx.doi.org/10.1016/j.rser.2016.01.055>

- [6] ALADE, Ibrahim Olanrewaju; RAHMAN, Mohd Amiruddin Abd; SALEH, Tawfik A. Predicting the specific heat capacity of alumina/ethylene glycol nanofluids using support vector regression model optimized with Bayesian algorithm. *Solar Energy*, [s.l.], v. 183, p.74-82, maio 2019. Elsevier BV. <http://dx.doi.org/10.1016/j.solener.2019.02.060>.
- [7] MICHAEL, Jee Joe; INIYAN, S. Performance of copper oxide/water nanofluid in a flat plate solar water heater under natural and forced circulations. *Energy Conversion And Management*, [s.l.], v. 95, p.160-169, maio 2015. Elsevier BV. <http://dx.doi.org/10.1016/j.enconman.2015.02.017>.
- [8] JIN, Haichuan et al. Photothermal conversion efficiency of nanofluids: An experimental and numerical study. *Solar Energy*, [s.l.], v. 139, p.278-289, dez. 2016. Elsevier BV. <http://dx.doi.org/10.1016/j.solener.2016.09.021>.
- [9] SHARAFELDIN, M.a.; GRÓF, Gyula. Evacuated tube solar collector performance using CeO₂/water nanofluid. *Journal Of Cleaner Production*, [s.l.], v. 185, p.347-356, jun. 2018. Elsevier BV. <http://dx.doi.org/10.1016/j.jclepro.2018.03.054>.
- [10] LIU, Qibin et al. Modeling and optimizing parabolic trough solar collector systems using the least squares support vector machine method. *Solar Energy*, [s.l.], v. 86, n. 7, p.1973-1980, jul. 2012. Elsevier BV. <http://dx.doi.org/10.1016/j.solener.2012.01.026>.
- [11] Cyril Voyant, Gilles Notton, Soteris Kalogirou, Marie-Laure Nivet, Christophe Paoli, Fabrice Motte, Alexis Foulloy, Machine learning methods for solar radiation forecasting: A review, *Renew. Energy* 105 (2017) 569-582, <https://doi.org/10.1016/j.renene.2016.12.095>.
- [12] FANG, Tingting; LAHDELMA, Risto. Evaluation of a multiple linear regression model and SARIMA model in forecasting heat demand for district heating system. *Applied Energy*, [s.l.], v. 179, p.544-552, out. 2016. Elsevier BV. <http://dx.doi.org/10.1016/j.apenergy.2016.06.133>.
- [13] JAMES, Gareth et al. *An Introduction to Statistical Learning: with Applications in R*. New York: Springer, 2013.
- [14] YANG, Xiaoxing; WEN, Wushao. Ridge and Lasso Regression Models for Cross-Version Defect Prediction. *Ieee Transactions On Reliability*, [s.l.], v. 67, n. 3, p.885-896, set. 2018. Institute of Electrical and Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/tr.2018.2847353>.
- [15] HE, Yaoyao et al. Electricity consumption probability density forecasting method based on LASSO-Quantile Regression Neural Network. *Applied Energy*, [s.l.], v. 233-234, p.565-575, jan. 2019. Elsevier BV. <http://dx.doi.org/10.1016/j.apenergy.2018.10.061>.
- [16] AHMAD, Muhammad Waseem; REYNOLDS, Jonathan; REZGUI, Yacine. Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees. *Journal Of Cleaner Production*, [s.l.], v. 203, p.810-821, dez. 2018. Elsevier BV. <http://dx.doi.org/10.1016/j.jclepro.2018.08.207>.
- [17] ALFADDA, Abdullah; RAHMAN, Saifur; PIPATTANASOMPORN, Manisa. Solar irradiance forecast using aerosols measurements: A data driven approach. *Solar Energy*, [s.l.], v. 170, p.924-939, ago. 2018. Elsevier BV. <http://dx.doi.org/10.1016/j.solener.2018.05.089>.
- [18] WANG, Shengzheng et al. Predicting ship fuel consumption based on LASSO regression. *Transportation Research Part D: Transport and Environment*, [s.l.], v. 65, p.817-824, dez. 2018. Elsevier BV. <http://dx.doi.org/10.1016/j.trd.2017.09.014>.
- [19] CHU, Yinghao; COIMBRA, Carlos F.m.. Short-term probabilistic forecasts for Direct Normal Irradiance. *Renewable Energy*, [s.l.], v. 101, p.526-536, fev. 2017. Elsevier BV. <http://dx.doi.org/10.1016/j.renene.2016.09.012>.
- [20] NONNENMACHER, Lukas; KAUR, Amanpreet; COIMBRA, Carlos F.m.. Day-ahead resource forecasting for concentrated solar power integration. *Renewable Energy*, [s.l.], v. 86, p.866-876, fev. 2016. Elsevier BV. <http://dx.doi.org/10.1016/j.renene.2015.08.068>.
- [21] AL-OBEIDAT, Feras; SPENCER, Bruce; ALFANDI, Omar. Consistently accurate forecasts of temperature within buildings from sensor data using ridge and lasso regression. *Future Generation Computer Systems*, [s.l.], abr. 2018. Elsevier BV. <http://dx.doi.org/10.1016/j.future.2018.02.035>