



Enhanced Support Vector Regression for Accurate CO₂ Emission Prediction from Energy Consumption Data

Oluwatayomi Rereloluwa Adegboye

University of Mediterranean Karpasia, Mersin-10, Northern Cyprus, TR-10 Mersin,
Turkey

aoluwatayomi@gmail.com

Abstract

Accurate CO₂ Emission prediction is essential for evidence-based climate policy, yet Support Vector Regression (SVR), a leading machine learning approach for prediction tasks, requires precise hyperparameter selection, which standard optimizers often address poorly due to premature convergence. This chapter proposes SCMSSA, a three-mechanism enhancement of the Salp Swarm Algorithm (SSA) that integrates logistic-chaotic single-dimensional perturbation, convex lens mirror imaging opposition, and sine-cosine follower perturbation to simultaneously improve exploration coverage, population diversity, and convergence speed, thereby optimizing the parameters of SVR. Applied to monthly U.S. Energy Information Administration records (1973-2023), the SVR-SCMSSA model attains an accuracy of 95.78%, surpassing all compared models on every evaluation metric.

Keywords: Salp Swarm Algorithm (SSA), Support Vector Regression (SVR), Energy Consumption, United States, Carbon Emission

1 Introduction

Fossil fuel combustion accounts for roughly 80% of global primary energy use and is the dominant anthropogenic driver of rising CO₂ concentrations (Calvin et al., 2023; Holechek et al., 2022; Luo et al., 2020). The Copenhagen Accord mandates limiting warming to 2°C requires a significant reduction in emissions relative to 1990 by 2050 yet concentrations continue to climb (Ramanathan & Xu, 2010; Schmalensee et al., 1998). Reliable CO₂ forecasting is therefore foundational for evaluating mitigation pathways and informing energy policy (Madejski et al., 2022; Yu & Du, 2019).

Machine learning hybrid architectures coupling regression models with metaheuristic optimizers have gained traction for environmental forecasting (Ahmed et al., 2021; Masood



& Ahmad, 2021; Qin et al., 2023). SVR (Cortes & Vapnik, 1995) is a strong candidate owing to its structural risk minimization and kernel-based nonlinear mapping, but its accuracy depends heavily on the penalty C and kernel γ (Malik et al., 2020) parameters, which are non-trivial to tune automatically. The SSA (Mirjalili et al., 2017) is a competitive optimizer valued for simplicity, yet its canonical form suffers from premature convergence and insufficient population diversity on multimodal landscapes (Si et al., 2022), limiting accuracy. The No-Free-Lunch theorem (Wolpert & Macready, 1997) motivates continued algorithm innovation; this work introduces SCMSSA, three complementary enhancements to SSA, and couples it with SVR for high-accuracy CO₂ prediction.

Related Work

Metaheuristic-Optimized ML for CO₂ and Environmental Prediction

The fusion of metaheuristic optimizers with ML regressors has produced competitive CO₂ and greenhouse gas (GHG) forecasting models. Ahmed et al. (Ahmed et al., 2021) coupled a Multi-Layer Perceptron with Nuclear Reaction Optimization for monthly streamflow prediction, demonstrating rapid convergence on long hydrometeorological records. For CO₂ forecasting specifically, Qin et al. (Qin et al., 2023) introduced an enhanced Marine Predator Algorithm paired with multi-kernel SVR to project China's emissions. Jalaee et al. (Jalaee et al., 2021) merged an ANN with a Cuckoo Optimization Algorithm (CO-ANN), yielding a reliable global CO₂ projections to 2050, while Uzlu (Uzlu, 2021) applied a Grey Wolf Optimizer-trained ANN to Turkey's GHG emissions, reducing relative error compared to alternative ANN optimizers. Each of these contributions relies on a single-mechanism optimizer that may underperform on rugged hyperparameter landscapes, motivating the multi-mechanism design of SCMSSA.

SSA Enhancements and Opposition-Based Learning

The canonical SSA (Faris et al., 2020; Mirjalili et al., 2017) has been widely recognized for its simplicity, yet practical applications confirm its susceptibility to local optima on multimodal problems (Mahapatra et al., 2023; Si et al., 2022). Si et al. 2022 addressed this by incorporating opposition-based learning (OBL) into SSA, significantly improving diversity and coverage. Independently, chaotic maps have been integrated into swarm optimizers by Luo et al. (Luo et al., 2020) in firework algorithms and Zhang et al. (Zhang



et al., 2022) in butterfly optimization, to inject ergodic perturbations that escape local optima. Tizhoosh (Tizhoosh, 2005) established OBL as a general machine-intelligence paradigm; Yao et al. (Yao et al., 2023) operationalized convex-lens imaging within the Enhanced Snake Optimizer (ESO), confirming geometry-based opposition as an effective acceleration strategy. No prior work unifies all three mechanisms-chaotic perturbation, lens-mirror imaging, and sine-cosine perturbation-within SSA, the gap addressed by SCMSSA.

Methodology

SSA and SVR Foundations

The SSA leader updates its position in dimension j as expressed in Eq. (1)

$$x_j^1 = \begin{cases} F_j + r_1 ((ub_j - lb_j)r_2 + lb_j) & r_3 \geq 0.5 \\ F_j - r_1 ((ub_j - lb_j)r_2 + lb_j) & r_3 < 0.5 \end{cases} \quad (1)$$

where $r_1 = 2e^{-(2t/T)^2}$ decays from exploration to exploitation (Faris et al., 2020; Mirjalili et al., 2017). SVR (Cortes & Vapnik, 1995) minimizes $\frac{1}{2} \|w\|^2 + C \sum_i (\varepsilon_i + \varepsilon_i^*)$ subject to ε -insensitive constraints, with the RBF kernel $K(x, x_i) = e^{-\gamma \|x - x_i\|^2}$. SCMSSA automates joint (C, γ) selection using RMSE as the fitness function.

SCMSSA Enhancement Mechanisms

Logistic Chaotic Single-Dimensional Perturbation (LCSDP). A logistic map $C(t+1) = \mu C(t)(1 - C(t))$ with $\mu = 4$ generates ergodic, non-repeating sequences (Li et al., 2020; Luo et al., 2020; Zhang et al., 2022). At each iteration, one dimension d of X_{best} is perturbed: $X_{\text{new},d} = lb_d + C(t)(ub_d - lb_d)$. The update is accepted only if fitness improves. Perturbing a single dimension preserves valuable information in the remaining coordinates while providing targeted disruption.

Convex Lens Mirror Imaging. Inspired by Opposition-Based Learning (Tizhoosh, 2005) and the lens-imaging mechanism of Yao et al. (Yao et al., 2023), two opposing candidates, a convex-image salp s^* and a plane-mirror salp s^{**} are generated for each individual. A nonlinear scaling factor $\delta(t) = 10 - (10 - 9)(t/T)^2$ governs lens thickness, favouring



Sine-Cosine Position Perturbation. Perturbation factors $p_{\sin}$ and $p_{\cos}$ functions of the iteration ratio t/T and the fitness gap ($f_{\text{best}}^* - f_i$) replace the canonical SSA follower update for salps $i > N/2$. Oscillatory disruption diversifies search trajectories and aids escape from local optima (Yao et al., 2023)

Experimental Results

CO₂ Prediction Performance

The U.S. EIA dataset comprises monthly observations (1973-2023) with eight variables: CO₂ (target), Fossil Fuels, Nuclear, Hydro, Geothermal, Wood, Waste, and Biomass. Features are min-max normalized to [0,1] and partitioned 70%/30% for training/testing. Tables 1 and 2 compare SVR-SCMSSA against unoptimized SVR and Extreme Learning Machine (ELM) on five metrics. SVR-SCMSSA leads all competitors on every metric for both splits. On the test set, RMSE decreases significantly compared to SVR, while the accuracy of 0.9578 exceeds the next-best ELM of 0.9286. ELM achieves training accuracy of 0.9477, but degrades in the test, indicating overfitting.

Table 1. Training results

Model	RMSE	R^2	MSE	MAE	MAPE
SVR-SCMSSA	0.00079	0.9586	0.00055	0.01683	0.2786
SVR	0.00084	0.9159	0.00102	0.01951	0.3227
ELM	0.04365	0.9477	0.00191	0.03212	0.7033

Table 2. Test results

Model	RMSE	R^2	MSE	MAE	MAPE
SVR-SCMSSA	0.00056	0.9578	0.00056	0.01634	0.2699
SVR	0.00133	0.8998	0.00133	0.03154	0.5200
ELM	0.04934	0.9286	0.00243	0.03520	0.2766



Conclusion

This chapter introduced SCMSSA, a three-mechanism enhancement of SSA that combines logistic-chaotic perturbation, convex-lens mirror imaging, and sine-cosine position perturbation. The SVR-SCMSSA hybrid applied to EIA monthly records outperformed all compared models across all metrics, achieving test accuracy of 0.9578 and an RMSE of 0.00056, confirming both high explanatory fidelity and tight error containment. These results are particularly notable given the inherent volatility and seasonal irregularity embedded in monthly industrial activity data, where conventional single-stage models often fail to maintain consistent generalization.

References

- Ahmed, A. N., Van Lam, T., Hung, N. D., Van Thieu, N., Kisi, O., & El-Shafie, A. (2021). A comprehensive comparison of recent developed meta-heuristic algorithms for streamflow time series forecasting problem. *Applied Soft Computing*, 105, 107282. <https://doi.org/10.1016/j.asoc.2021.107282>
- Calvin, K., Dasgupta, D., Krinner, G., Mukherji, A., Thorne, P. W., Trisos, C., Romero, J., Aldunce, P., Barrett, K., Blanco, G., Cheung, W. W. L., Connors, S., Denton, F., Diongue-Niang, A., Dodman, D., Garschagen, M., Geden, O., Hayward, B., Jones, C., ... Péan, C. (with Lee, H.). (2023). *IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]*. IPCC, Geneva, Switzerland. (First). Intergovernmental Panel on Climate Change (IPCC). <https://doi.org/10.59327/IPCC/AR6-9789291691647>
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297. <https://doi.org/10.1007/BF00994018>
- Faris, H., Mirjalili, S., Aljarah, I., Mafarja, M., & Heidari, A. A. (2020). Salp Swarm Algorithm: Theory, Literature Review, and Application in Extreme Learning Machines. In S. Mirjalili, J. Song Dong, & A. Lewis (Eds.), *Nature-Inspired Optimizers: Theories, Literature Reviews and Applications* (pp. 185–199). Springer International Publishing. https://doi.org/10.1007/978-3-030-12127-3_11
- Holechek, J. L., Geli, H. M. E., Sawalhah, M. N., & Valdez, R. (2022). A Global Assessment: Can Renewable Energy Replace Fossil Fuels by 2050? *Sustainability*, 14(8), 4792. <https://doi.org/10.3390/su14084792>
- Jalaei, S. A., Shakibaei, A., Akbarifard, H., Horry, H. R., GhasemiNejad, A., Nazari Robati, F., Amani zarin, N., & Derakhshani, R. (2021). A novel hybrid method



- Li, C., Li, J., & Chen, H. (2020). A Meta-Heuristic-Based Approach for Qos-Aware Service Composition. *IEEE Access*, 8, 69579–69592. IEEE Access. <https://doi.org/10.1109/ACCESS.2020.2987078>
- Luo, W., Jin, H., Li, H., Fang, X., & Zhou, R. (2020). Optimal Performance and Application for Firework Algorithm Using a Novel Chaotic Approach. *IEEE Access*, 8, 120798–120817. <https://doi.org/10.1109/ACCESS.2020.3004430>
- Madejski, P., Chmiel, K., Subramanian, N., & Kuś, T. (2022). Methods and Techniques for CO₂ Capture: Review of Potential Solutions and Applications in Modern Energy Technologies. *Energies*, 15(3), 887. <https://doi.org/10.3390/en15030887>
- Mahapatra, A. K., Panda, N., & Pattanayak, B. K. (2023). Quantized Salp Swarm Algorithm (QSSA) for optimal feature selection. *International Journal of Information Technology*, 15(2), 725–734. <https://doi.org/10.1007/s41870-023-01161-6>
- Malik, A., Tikhamarine, Y., Souag-Gamane, D., Kisi, O., & Pham, Q. B. (2020). Support vector regression optimized by meta-heuristic algorithms for daily streamflow prediction. *Stochastic Environmental Research and Risk Assessment*, 34(11), 1755–1773. <https://doi.org/10.1007/s00477-020-01874-1>
- Masood, A., & Ahmad, K. (2021). A review on emerging artificial intelligence (AI) techniques for air pollution forecasting: Fundamentals, application and performance. *Journal of Cleaner Production*, 322, 129072. <https://doi.org/10.1016/j.jclepro.2021.129072>
- Mirjalili, S., Gandomi, A. H., Mirjalili, S. Z., Saremi, S., Faris, H., & Mirjalili, S. M. (2017). Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems. *Advances in Engineering Software*, 114, 163–191. <https://doi.org/10.1016/j.advengsoft.2017.07.002>
- Qin, X., Zhang, S., Dong, X., Zhan, Y., Wang, R., & Xu, D. (2023). China's carbon dioxide emission forecast based on improved marine predator algorithm and multi-kernel support vector regression. *Environmental Science and Pollution Research*, 30(3), 5730–5748. <https://doi.org/10.1007/s11356-022-22302-7>
- Ramanathan, V., & Xu, Y. (2010). The Copenhagen Accord for limiting global warming: Criteria, constraints, and available avenues. *Proceedings of the National Academy of Sciences*, 107(18), 8055–8062. <https://doi.org/10.1073/pnas.1002293107>
- Schmalensee, R., Stoker, T. M., & Judson, R. A. (1998). World Carbon Dioxide Emissions: 1950–2050. *The Review of Economics and Statistics*, 80(1), 15–27.
- Si, T., Miranda, P. B. C., & Bhattacharya, D. (2022). Novel enhanced Salp Swarm Algorithms using opposition-based learning schemes for global optimization problems. *Expert Systems with Applications*, 207, 117961. <https://doi.org/10.1016/j.eswa.2022.117961>
- Tizhoosh, H. R. (2005). Opposition-Based Learning: A New Scheme for Machine Intelligence. *International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06)*, 1, 695–701. <https://doi.org/10.1109/CIMCA.2005.1631345>



-
- Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, 1(1), 67–82. IEEE Transactions on Evolutionary Computation. <https://doi.org/10.1109/4235.585893>
- Yao, L., Yuan, P., Tsai, C.-Y., Zhang, T., Lu, Y., & Ding, S. (2023). ESO: An enhanced snake optimizer for real-world engineering problems. *Expert Systems with Applications*, 230, 120594. <https://doi.org/10.1016/j.eswa.2023.120594>
- Yu, Y., & Du, Y. (2019). Impact of technological innovation on CO2 emissions and emissions trend prediction on ‘New Normal’ economy in China. *Atmospheric Pollution Research*, 10(1), 152–161. <https://doi.org/10.1016/j.apr.2018.07.005>
- Zhang, M., Wang, D., & Yang, J. (2022). Hybrid-Flash Butterfly Optimization Algorithm with Logistic Mapping for Solving the Engineering Constrained Optimization Problems. *Entropy*, 24(4), Article 4. <https://doi.org/10.3390/e24040525>