UDC 004.85:004.05
DOI https://doi.org/10.32782/tnv-tech.2024.6.12

MULTI-OUTPUT REGRESSION MODELS FOR CONTROLLING MULTICOMPONENT DYNAMIC SYSTEMS

Symonov D. I. *– PhD, Research Assistant of the V. M. Glushkov Institute of Cybernetics of the National Academy of Sciences of Ukraine*
ORCID ID: 0000-0002-6648-4736

Zaika B. Yu. *– Postgraduate Student of the V. M. Glushkov Institute of Cybernetics of the National Academy of Sciences of Ukraine*
ORCID ID: 0009-0001-9567-8361

Symonov Ye. D. *– Master of the V. M. Glushkov Institute of Cybernetics of the National Academy of Sciences of Ukraine*
ORCID ID: 0009-0008-2581-2001

Modern multi-component systems are characterized by the interaction of numerous internal components and external factors, which can exhibit both regular and chaotic behavior. Effective management of such systems requires tools capable of providing accurate state predictions under conditions of uncertainty and limited input data. This article explores the use of multi-output regression models, which enable the consideration of interdependencies among system components, optimization of the parametric space, and improvement in prediction accuracy. Multi-output models allow simultaneous forecasting of several aspects of a system's state, reducing errors and enhancing the generalization ability of the models. The article provides a detailed examination of methods to improve such models, including minimization of noise influence, accounting for the temporal scales of component changes, optimization for small data samples, and increasing the interpretability of predictions. Approaches to addressing data scarcity are proposed, such as knowledge sharing between tasks and the use of generative models. Special attention is given to the challenges of applying multi-output models, including the risks of overfitting, conflicts between optimization objectives, and the impact of correlation biases. Strategies to mitigate these risks are discussed, including adapting multi-criteria optimization, parameter regularization, and developing hierarchical models that can account for system dynamics across different time scales. Ensemble approaches, which integrate the outputs of sub-models into a unified architecture, are highlighted for their ability to enhance noise robustness, prediction accuracy, and model adaptability to changing conditions. The approaches proposed in the article have practical significance for automating decision-making processes in complex multi-component systems operating under high variability and data limitations. This provides a comprehensive framework for forecasting, contributing to more effective management of dynamic systems across various domains. Thus, the article makes a significant contribution to the development of methodologies for modeling complex systems, expanding the possibilities for their analysis and management.

Key words: *multivariate regression models, multicomponent systems, system state prediction, ensemble approaches, regularization.*

Симонов Д. І., Заіка Б. Ю., Симонов Є. Д. Мультивихідні регресійні моделі для управління багатокомпонентними динамічними системами

Сучасні багатокомпонентні системи визначаються взаємодією численних внутрішніх компонентів і зовнішніх факторів, які можуть мати як регулярний, так і хаотичний характер. Ефективне управління такими системами вимагає інструментів, здатних забезпечувати точне прогнозування стану за умов невизначеності та обмеженості вхідних даних. У статті досліджено використання мультивихідних регресійних моделей, що дозволяють враховувати взаємозалежності між компонентами системи, оптимізовувати параметричний простір і підвищувати точність прогнозування. Мультивихідні моделі забезпечують одночасне прогнозування кількох аспектів стану системи, знижуючи похибки та підвищуючи узагальнювальну здатність моделей. У статті детально розглянуто методи вдосконалення таких моделей, серед яких мінімізація впливу шуму, *врахування часових масштабів змін компонентів, оптимізація для малих вибірок даних, а також підвищення інтерпретованості прогнозів. Запропоновано підходи до роботи з малою кількістю даних, включаючи обмін знаннями між задачами та використання генеративних моделей. Особливу увагу надано викликам, які виникають при застосуванні мультивихідних моделей, зокрема ризику перенавчання, конфліктам між цілями оптимізації та впливу кореляційних упереджень. Розглянуто способи мінімізації цих ризиків, зокрема адаптацію багатокритеріальної оптимізації, регуляризацію параметрів, а також розробку ієрархічних моделей, здатних враховувати динаміку систем на різних часових рівнях. Виділяються ансамблеві підходи, які дозволяють інтегрувати результати підмоделей у єдину архітектуру для підвищення стійкості до шуму, точності прогнозування та адаптивності моделей до змінних умов. Запропоновані у статті підходи мають практичну значущість для автоматизації процесу прийняття рішень у складних багатокомпонентних системах, що функціонують в умовах високої варіативності та обмеженості даних. Це забезпечує комплексний підхід до прогнозування, що сприяє ефективнішому управлінню динамічними системами у різних галузях. Таким чином, стаття робить значний внесок у розвиток методології моделювання складних систем і розширює можливості їх аналізу та управління.*

Ключові слова: *мультивихідні регресійні моделі, багатокомпонентні системи, прогнозування стану систем, ансамблеві підходи, регуляризація.*

Introduction. Modern multi-component systems are characterised by a high level of complexity, dependence on numerous internal and external factors that can be both regular and chaotic. The study of such systems requires the use of effective modelling methods that can take into account the interaction of components and predict the behaviour of the system under uncertainty and limited data [1, 2].

One of the most promising approaches is the use of multi-output regression models that provide simultaneous forecasting of several aspects of the system state. These models allow taking into account correlations between components and aspects of the state, reduce the parameter space and optimise loss functions, increasing the accuracy and reliability of forecasts. However, the implementation of such models is complicated by certain issues, including the presence of anomalies and noise in the data, limited size of training samples, a variety of temporal changes in components, and the complexity of interpreting the results [3–5].

Overcoming these problems requires the introduction of innovative approaches to minimise risks, optimise resources and ensure model stability. Particular attention should be paid to methods of working with small samples, the use of common parameters, knowledge transfer and generative models. This helps to improve the accuracy and adaptability of models to uncertainty, which is critical for the management, forecasting and optimisation of multi-component systems.

The relevance of the study is due to the growing need for modelling complex systems in such fields as engineering, economics and medicine. Traditional approaches to modelling are often ineffective due to the neglect of these aspects. The approaches proposed in this article not only improve the accuracy of forecasts, but also provide greater flexibility and adaptability of models. The study of this topic has significant scientific and practical potential, contributing to the development of a methodology for modelling complex systems and algorithms for effective management of multicomponent systems in modern conditions.

Problem statement. Multicomponent systems are complex objects, the dynamics of which largely depends on the interaction of internal components $X(t)=\{x_i(t)\}$ and the influence of external factors $U(t)$. Prediction of the state of such systems is complicated by a high level of noise $\Xi(t)$, different rates of changes in the state of components, limited data sampling of certain aspects of the system state y_i, which negatively affects the quality and ability to interpret the results of modelling (forecasting), respectively, reducing the likelihood of practical implementation of these models to automate the decision-making process on the functioning of multicomponent systems in real conditions.

The aim of the study. The purpose of this study is to develop and improve methods for minimising risks and increasing the accuracy of multi-output regression models for predicting the state of multi-component systems.

Analysis of recent research and publications. Recent research in the field of forecasting the state of multicomponent systems has paid considerable attention to the use of multi-output regression models that allow for the consideration of interdependencies between aspects of the system state. In particular, works [4] and [6] demonstrate the effectiveness of such models in problems with limited data, suggesting the use of common parameters and multitasking optimisation to improve the accuracy of forecasts. In addition, study [5] emphasises the importance of regularising model parameters to reduce the risk of overfitting and improve generalization.

The approaches to noise and uncertainty accounting, which are discussed in publications [7] and [8], are of considerable interest. These papers propose noise filtering methods, such as the Kalman filter, as well as adaptive regularisation that takes into account the signal-to-noise ratio (SNR) in the system components. The research results confirm that these methods provide a significant reduction in the impact of noise components and improve the accuracy of forecasting in conditions of high data variability.

Particular attention is drawn to the works devoted to modelling the dynamics of systems with different time scales. For example, in [9], a network with heterogeneous leaky integrator neurons was proposed to efficiently model and predict multiscale dynamics by adaptively selecting time scales during the training process. Other researchers, such as [10], use wavelet transform methods to extract time trends, which allows taking into account complex system dynamics at different time levels.

The problem of model interpretability is highlighted in studies [11–13], which focus on the use of SHAP (SHapley Additive Explanations) and LIME (Local Interpretable Model Agnostic Explanations) methods to analyse the relationships between system components. These approaches help to ensure the transparency of forecasts and increase the credibility of models in practical applications.

Despite significant progress in research, a number of challenges remain unresolved, including the integration of approaches to working with small samples, ensuring the adaptability of models to noise and time scales, and developing interpretability in highly complex systems. This determines the need for further research aimed at improving existing methods and developing new approaches to modelling multicomponent systems.

Presentation of the main research material. As noted above, the state of a multicomponent system can largely depend on external and internal chaotic factors that affect its dynamics. These factors may include irregular changes in input parameters, external factors, and internal noise inherent in many real systems.

The state formula of a multicomponent system $Y(t)$ can be represented in a general form through the state vector of system aspects that describe all important parameters of the multicomponent system [2, 15]:

$$Y(t)=\left[y_1(t),y_2(t),...,y_m(t),\omega_j,t\right]^T=\begin{bmatrix} f_1\left(x_i(t),\omega_i,t\right) \\ f_2\left(x_i(t),\omega_i,t\right) \\ \dots \\ f_m\left(x_i(t),\omega_i,t\right) \end{bmatrix}, \tag{1}$$

where $x_i(t)$ is the vector of the state of the system components at time *t*;

ω_i is the weighting coefficients of the system components $x_i(t)$;

ω_j is the weighting coefficients of the system aspects $y_i(t)$.

The state vector of a system component $x_i(t)$ at time *t* can be defined as a system of iterative equations:

$$x_i(t+1) = f_i\left(x_i(t), \sum_{j=1}^{n} A_{ij} \cdot g(x_j(t), t), U(t), \Xi(t)\right), \tag{2}$$

where f is a function describing the evolution of the *i*-th component of the system;

A_{ij} is an element of the influence matrix between components *i* and *j*;

$g(x_j(t), t)$ is a function describing the influence of component *j* on *i*;

$U(t)$ is a vector of external factors that affect the state of the system;

$\Xi(t)$ is a vector of noise parameters for all components of the system, $\Xi(t) = [\xi_1(t), \xi_2(t), ..., \xi_n(t)]^T$;

$\xi_i(t)$ is a random variable.

Since it is intended to predict the state of a dynamic multicomponent system, it is advisable to consider a random variable $\xi_i(t)$ taking into account time series statistics [16]:

$$\xi_i(t) = \sigma_i(t) \cdot \varsigma_i(t), \tag{3}$$

where $\sigma_i(t)$ is the time modulation of the intensity of changes in the state of the system component; $\varsigma_i(t) \sim \mathcal{N}(0,1)$.

Equations (1) and (2) reflect not only the state of individual components, but also the differences between phase points, which complicates the analysis and forecasting of the system. Even a slight discrepancy in the input data can significantly affect the modelling results, especially if the system has lost its equilibrium state. One of the approaches to improve the quality of forecasting the state of multi-component systems is the multi-output approach.

The advantages of using the multi-output approach include the following aspects:

1. *Taking into account interdependencies between state aspects*: a multi-output approach to predicting the state of a multicomponent system allows simultaneously taking into account interdependencies between different state aspects, which increases the accuracy and generalization of the model. For example, if there is a correlation between aspects, the regression function $f_i(\cdot)$ is optimised simultaneously for all output variables, which reduces the probability of error compared to independent forecasting of each aspect y_i.

If the aspects of the state $y_1, y_2, ..., y_m$ have a correlation $\rho(y_i, y_j)$, $i = \overline{1,m}, j = \overline{1,m}, i \neq j$, then joint forecasting within the framework of a multi-output model allows taking into account the mutual influence of the aspects and the components $x_i(t)$, through the model parameters, which positively affects the value of the mean square error (MSE) of the forecast:

$$MSE_{mult} = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2 \leq \sum_{i=1}^{m} MSE_i, \tag{4}$$

where $\hat{y}_i$ is the predicted value of y_i.

2. *Reducing dimensionality and improving generalization*: multi-output models use a single architecture to predict all aspects of the state, which allows for a reduction in the number of parameters compared to sequential (independent) models, meaning that the parameter space *d* can be reduced:

$$d_{mult} = d_{shared} + d_{spec} \ll \sum_{i=1}^{m} d_i, \tag{5}$$

where d_{shared} – the parameters that are similar for all aspects y_i;

d_{spec} – the parameters that take into account the individual characteristics of each aspect y_i;

d_i – the model parameters for predicting each aspect y_i.

3. *Improving prediction with limited data:* in problems with a few observations for each aspect y_i, a multi-output model allows knowledge to be transferred between aspects y_i using common parameters $x_i(t)$. This property is especially important when the state aspects share a similar nature or have similar patterns. In this case, the joint loss function is minimised:

$$L(\Theta) = \sum_{i=1}^{m} L_i(\Theta, X, Y), \tag{6}$$

where $L_i(\Theta, X, Y)$ is the loss function for the *i*-th aspect;

Θ is a set of common model parameters that are optimised during training.

This allows aspects y_i with less data to gain useful information from other aspects $y_j, i \neq j$, by optimising shared parameters Θ.

4. *Support for scenario analysis and decision-making*: a multi-output model allows simulating the behaviour of the system in different scenarios simultaneously, evaluating the impact of changes in parameters or external factors on several aspects of the state y_i. This approach increases the usefulness of the model in the tasks of managing a multi-component system. For example, if a system with two state aspects y_i (e.g., equilibrium state and system stability) is analysed, a multi-output model can evaluate the trade-off between these aspects when input parameters change [16, 17].

Thus, the use of a multi-output regression model to predict the state of a multicomponent system allows for an integrated approach to forecasting, reducing the forecasting error, decreasing the dimensionality of the parameter space, and reducing the requirements for the size of the input data sample without losing the quality of forecasting.

Despite the numerous advantages of multi-output regression models, their use is associated with a number of potential risks. These risks can lead to model degradation, reduced efficiency and forecasting accuracy. The risks of using a multivariate approach include the following aspects:

1. *Model degradation due to conflicting optimisation objectives*: when simultaneously predicting several aspects of the model's state, several loss functions need to be optimised. If the state aspects have contradictory dependencies or different natures, the model may not generalise the data well enough. This occurs when minimising the loss function for one aspect worsens the forecast accuracy for another, meaning that the gradients of the loss functions $\nabla_{\Theta} L_i(\Theta)$ are oriented in different directions, thus $\nabla_{\Theta} L_i(\Theta) \cdot \nabla_{\Theta} L_j(\Theta) < 0, \exists y : i \neq j$. In such cases, parameter optimisation can lead to local minimums that do not provide high accuracy for all aspects.

2. *Excessive complexity of model training (overfitting)*: a multi-output model may have a significantly higher number of parameters compared to models that predict a single aspect. This creates a risk of overfitting, especially if the available data is limited. In such cases, the model demonstrates good accuracy on the training data, but poor generalisation on the test data [18].

3. *Vulnerability to correlation bias*: if there is a high correlation $\rho(y_i, y_j)$ between aspects of state y_i, the model may use these dependencies for prediction without taking into account the fundamental cause and effect relationships. This leads to the construction of models that show degradation when conditions or data distribution change.

4. *Failure to take into account different time scales*: if aspects of the state change with different time scales (e.g., some aspects have fast dynamics and others have slow

dynamics), the multi-output model may not process them correctly [9, 10]. For example, if fast changes are described by the temporal modulation of the intensity of changes in the state of a system component $\sigma_1(t)$, and slow changes are described by $\sigma_2(t)$, then modelling the entire system without taking these frequencies into account can lead to the loss of important information, since the model cannot learn optimally on different time scales: $Y(t+1)=F(X(t),\sigma_1,\sigma_2,t)$.

5. *The influence of the random component*: if the noise component $\Xi(t)$ differs significantly between aspects of the state y_i, the model may incorrectly estimate the weight of different aspects ω_i. This leads to a decrease in accuracy for aspects with a low signal-to-noise ratio (SNR) [3].

6. *Difficulty in interpreting results*: multi-output models can be difficult to interpret, especially if aspects of the state y_i interact through nonlinear dependencies. This makes it difficult to assess the contribution of individual system components $x_i(t)$ to the forecast $\hat{y}_i$. Without appropriate analysis tools, such as SHAP or LIME, it is difficult to verify the correctness of the results [11–13].

Solving these problems requires appropriate algorithmic and methodological approaches. Accordingly, to mitigate the risks associated with the use of multi-output regression models in predicting the state of a multicomponent system, appropriate methodological approaches and algorithmic strategies should be applied. The key recommendations for minimising these risks include the following:

1. *Resolving conflicting optimisation objectives*: To resolve conflicts between loss functions, multicriteria optimisation approaches need to be adapted. One method is to use dynamic weighting of loss functions:

$$\alpha_i(t)=\frac{\mathbb{E}[\nabla_\Theta L_i]}{\sum_{j=1}^{m}\mathbb{E}[\nabla_\Theta L_i]}, \tag{7}$$

where $\alpha_i(t)$ is the weighting factor for the *i*-th loss function, adapted depending on the size of the gradient.

This approach allows balancing the influence of loss functions in the learning process, especially when combined with methods of sequential parameter optimisation:

$$\Theta(t+1)=\Theta(t)-\eta\sum_{l=1}^{m}\omega_l(t)\cdot\nabla_\Theta L_i(\Theta(t)), \tag{8}$$

where $\omega_l(t)$ is a weighting factor that depends on the stage of training *l.*

2. *Preventing correlation bias*: To deal with correlation bias, it is necessary to separate correlations between aspects and causal relationships. For this purpose, it is advisable to use methods that introduce a penalty for high correlations between forecasts:

$$L_{decor}=L_i(\Theta(t))+\beta\sum_{i\neq j}\rho(\hat{y}_i,\hat{y}_j), \tag{9}$$

where ρ is the Pearson's correlation coefficient.

3. *Taking into account different time scales*: to model systems with different time scales, it is advisable to use hierarchical recurrent neural networks (HRNN) [19], which allow modelling dynamics at different levels:

$$h_i=h_{i,low}(t)+h_{i,higt}(t), \tag{10}$$

where $h_{i,low}(t)$ is the hidden state of the low-level block, $h_{i,low}(t)=f_{low}(h_{i,low}(t-1),X(t))$; $h_{i,higt}(t)$ is the hidden state of the high-level block, $h_{i,higt}(t)=f_{higt}(h_{i,higt}(t-1),h_{i,low}(t))$.

4. *Regularisation of model parameters*: one of the most effective ways to avoid overfitting is regularisation, which reduces the complexity of the model by limiting the values of its parameters. In the case of multi-output models, regularisation can be applied to the parameters Θ_s and Θ_i, where Θ_s represents the parameters that are common to all aspects y_i and Θ_i represents the parameters that are individual to the aspects y_i. The loss function with additional regularisation components has the form:

$$L_{add} = \sum_{i=1}^{n} \varpi_i \cdot L_i\left(\hat{y}_i, y_i\right) + \lambda_1 \left\|\Theta_s\right\|_2^2 + \lambda_2 \sum_{i=1}^{n} \left\|\Theta_i\right\|_2^2, \tag{11}$$

where λ_1, λ_2 are the hyperparameters of regularisation;

$\|\cdot\|_2^2$ is the ϖ_i-norm of the model parameters;

ϖ_i is the weighting factor, which implements the strategy of equality by the variance of the loss functions, $\varpi_i = \frac{1}{Var(L_i)}$.

This approach reduces the impact of small samples on model complexity and improves generalization.

5. *Noise regularisation*: to minimise the impact of noise on the model parameters Θ_s та Θ_i, it is advisable to use a specialised regularisation that takes into account the signal-to-noise ratio (SNR):

$$L_{SNR} = \sum_{i=1}^{n} L_i\left(\hat{y}_i, y_i\right) + \lambda \sum_{i=1}^{n} \frac{\sigma_\Xi^2}{\sigma_{X_i}^2}, \tag{12}$$

where σ_Ξ^2 – noise variance;

$\sigma_{X_i}^2$ – variance of the useful signal of the parameter $x_i(t)$ for the aspect y_i;

$$L_i\left(\hat{y}_i, y_i\right) = \begin{cases} \delta\left|\hat{y}_i - y_i\right| - \frac{\delta^2}{2}, \forall \Delta = \left|\hat{y}_i - y_i\right| > \delta, \\ \frac{1}{2}\left(\hat{y}_i - y_i\right)^2, \forall \Delta = \left|\hat{y}_i - y_i\right| \le \delta \end{cases},$$

where δ – threshold value for abnormal values.

6. *Transfer learning*: transferring knowledge from similar tasks or aspects can significantly improve the quality of the model. In this approach, the model is pre-trained on a large sample of data for one or more aspects of the state and then adapted for aspects with a small sample [20]. Accordingly, the adapted model for aspects with a small sample will have the form:

$$\Theta' = \arg\min_{\Theta'} \sum_{i=1}^{n} L_i\left(\hat{y}_i, y_i\right) + \lambda \sum_{i=1}^{n} \left\|\Theta - \Theta'\right\|_2^2, \tag{13}$$

where $\Theta = \arg\min_{\Theta} \sum_{j=1}^{m} L_j\left(\hat{y}_j, y_j\right)$ is the training of the base model.

7. *Reducing model complexity*: pruning is an effective approach to reduce model complexity [21], which minimises the risk of overfitting and increases generalisability, especially for multi-output regression models. The main idea of pruning is to remove unimportant parameters Θ_s and Θ_i, or system components $X_i(t)$ that have a low impact on the forecasting results.

In a model with parameters $\Theta = \left[\Theta_s, \Theta_i\right]$, the impact of each parameter on the loss function L_i is analysed. Parameters whose contribution to the reduction L_i is insignificant are removed:

$$\Theta' = \left\{ \Theta_j \in \Theta \left| \left|\nabla_{\Theta_j} L_i\left(\hat{y}_i, y_i\right)\right| > \varepsilon \wedge \frac{\sigma_\Xi^2}{\sigma_{X_i}^2} \le \eta \right.\right\}, \tag{14}$$

where ε is the threshold for determining the significance of the parameter;

η is the threshold for the noise to signal ratio.

Pruning is an appropriate method for improving the efficiency of multi-output regression models. Its application allows to reduce the complexity of the model, reduce the risk of overfitting and increase interpretability, while maintaining the accuracy of forecasting the state of a multicomponent system.

The generalisation of the considered methods of minimising risks and improving forecasting accuracy allows developing an ensemble of models that effectively takes into account the complexity of multicomponent systems. This approach is based on combining different methodologies into a single architecture to improve forecasting quality, noise immunity, and interpretability. Ensemble approaches based on a multi-output structure allow integrating the results of submodels into an ensemble average of forecasts:

$$\hat{y}_i = \frac{1}{\left|M_i\right|}\sum_{j\in M_i} M_i\left(X_i\left(t\right),\Theta_j\right), \tag{15}$$

where is the set of submodels that take into account the aspect y_i.

Integration of risk minimisation methods within the ensemble of models allows achieving high forecasting accuracy, reducing model complexity and increasing their resistance to noise factors. The proposed approach provides a comprehensive analysis of multicomponent systems and meets the modern requirements of scientific research aimed at modelling complex dynamic objects.

Results of an experimental study. To confirm the effectiveness of the proposed approaches related to the use of multi-output regression models in predicting the state of multicomponent dynamic systems, an experimental study was conducted on the open dataset "Energy Efficiency" [22]. The aim of the study was to empirically assess the accuracy, noise resistance, generalisability of the models and their interpretability in conditions of high data variability and different noise

Comparison of prediction accuracy: to demonstrate the improvement in prediction accuracy using the proposed methods, three multi-output models were trained and compared on the test set: Linear Regression, Ridge Regression (Linear Regression with regularization), and an ensemble model combining Linear Regression, Ridge Regression and RandomForestRegressor.

The Mean Squared Errors (MSEs) for two outputs are summarized in Table 1. The results indicate that applying regularization with Ridge Regression slightly reduces the prediction error compared to Linear Regression. Furthermore, the ensemble model achieves a significant reduction in error, showcasing its superior performance.

Figure 1 presents a plot comparing the predicted values to the real values. The ensemble model's predictions are the closest to the diagonal line, confirming that they align most closely with the real values. This observation further validates the effectiveness of the ensemble approach.

Table 1

MSEs of the models for both outputs

Model name	**MSE Y_1**	**MSE Y_2**
Linear Regression	9.7296	10.2100
Ridge Regression	9.2142	9.9373
Ensemble model	4.3177	5.5751

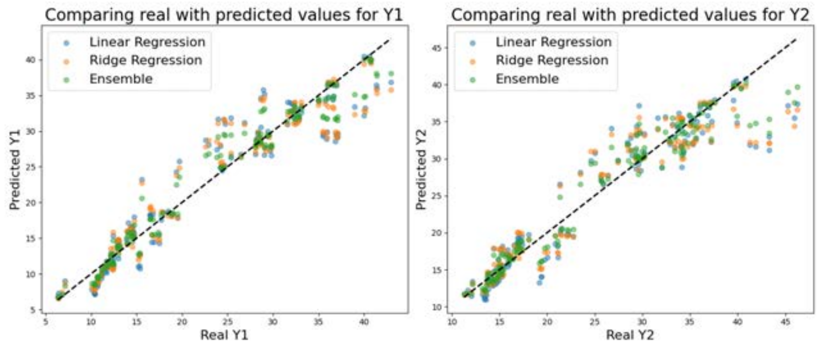

Fig. 1. Comparing the predicted values to the real

Noise vulnerability: To evaluate the resilience of the proposed approach to noise, the same models were trained on datasets with varying levels of added noise. The noise followed a normal distribution $N(0, \sigma^2)$, where σ ranged from 0 to 0.6. The average Mean Squared Errors (MSEs) for both outputs of the trained models are presented in Figure 2.

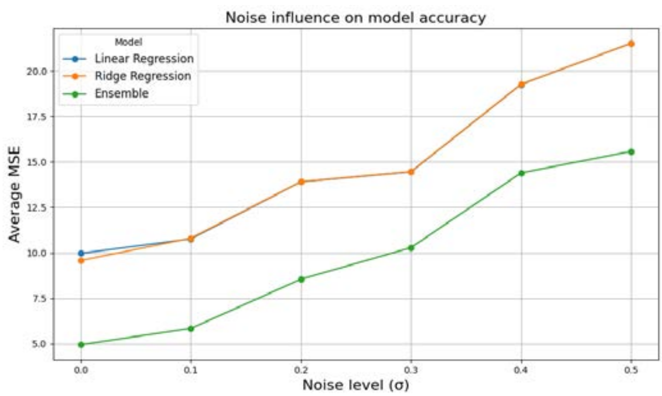

Fig. 2. Average MSE of models on datasets with varying noise level

The results demonstrate that, across all levels of noise, the ensemble approach consistently achieved the lowest average MSE. This highlights its robustness, even as the noise level increased.

Generalization ability: to assess the generalization ability of the proposed approach, the models were trained using varying portions of the dataset, ranging from 10% to 90% of the total data. The average Mean Squared Error (MSE) for each model was measured on test set and is presented in Figure 3.

The results show that the ensemble model consistently achieves the lowest average MSE across all training set sizes. While the performance of Linear Regression and Ridge Regression improves as the training data size increases, their error rates remain higher compared to the ensemble model. Notably, the ensemble model maintains its superiority even when trained on smaller portions of the dataset, demonstrating its ability to generalize effectively to unseen data.

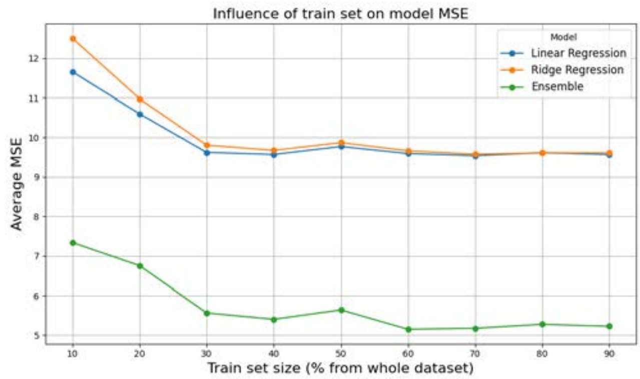

Fig. 3. Influence of train set on model MSE

These results indicate that the ensemble approach is not only more accurate but also more robust when training data is limited. This highlights its potential for achieving reliable predictions while mitigating the risk of overfitting, regardless of the amount of training data available.

Model interpretability. SHAP value plots were generated to evaluate the interpretability of the ensemble model and compare it with other multi-output models, including Linear Regression and Ridge Regression, for both outputs Y1 and Y2. The results are displayed in Figure 4 and allow us to understand the importance of features and their impact on model predictions.

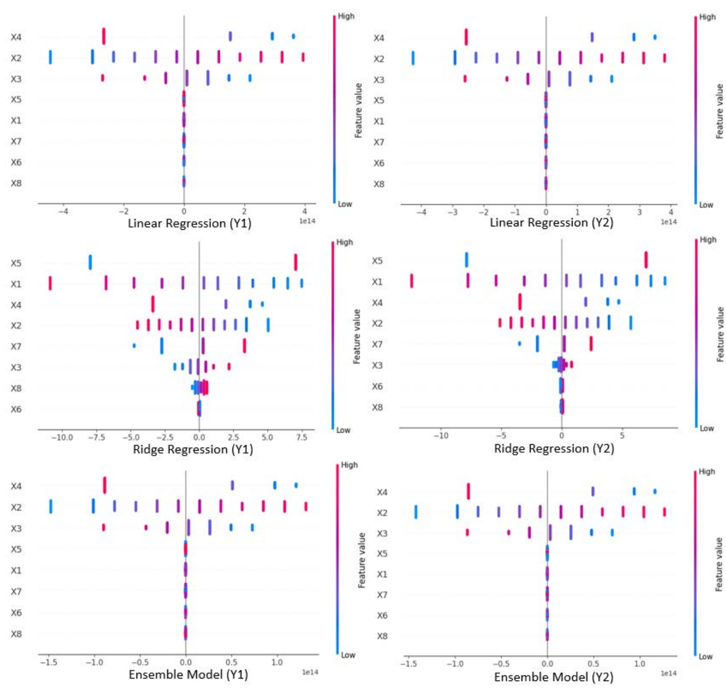

Fig. 4. SHAP values for each model and output

The SHAP value plots reveal that features X4, X2, and X1 are the most influential for predicting both outputs Y1 and Y2 across all models. Ridge Regression, compared to Linear Regression, reduces the magnitude of SHAP values, demonstrating that regularization effectively limits the dominance of individual features and enhances model stability.

Although the SHAP values for the ensemble model narrower around zero than those for Linear Regression, they exhibit a wider distribution of feature impacts compared to Ridge Regression. This emphasizes the role of regularization in improving stability. Despite its more complex structure, the ensemble model remains as interpretable with SHAP as the simpler models, making it both powerful and transparent.

Sensitivity to parameters: to evaluate the sensitivity of the ensemble model to its parameters, a heatmap was constructed to illustrate the influence of λ and depth on the model's Mean Squared Error (MSE), as shown in Figure 5.

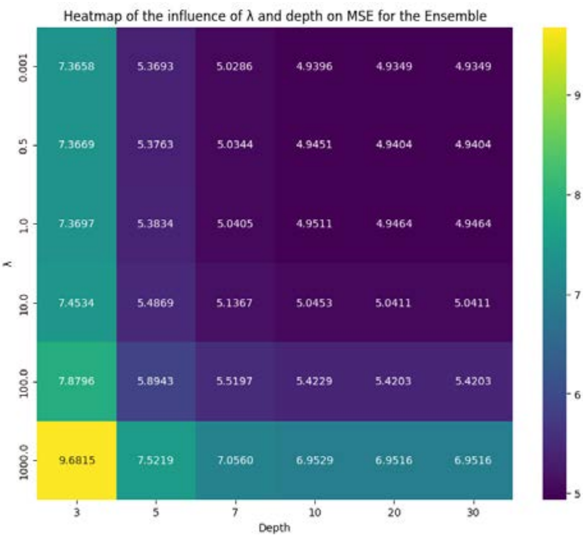

Fig. 5. Heatmap of the influence of λ and depth on MSE for the Ensemble model

The results demonstrate that the ensemble model maintains a high degree of stability across a wide range of λ and depth values. In most cases, the MSE remains consistently low, particularly as the depth increases, even when regularization strength (λ) varies significantly. While extreme values of λ (e.g., 1000) result in slightly higher MSE, the performance remains competitive and does not degrade below the levels observed previously in other models. Overall, the ensemble model achieves robust and stable performance, balancing regularization and complexity effectively. The observed stability at lower depths confirms that pruning the model is a useful strategy to reduce computational cost while preserving accuracy.

Conclusions. This study demonstrates the effectiveness of multi-output regression models for predicting the state of multi-component systems under conditions of limited data and high variability. The proposed risk minimization methods, such as multi-objective optimization, parameter regularization, and consideration of different time scales, contribute to improving model robustness to noise and enhancing prediction accuracy. The use of ensemble approaches allows for the integration of sub-model results, ensuring model adaptability to changing conditions and improving their generalization ability. Key challenges, such as conflicting optimization objectives, the risk of overfitting, and the complexity of interpretation, have been identified and require further research and refinement. The presented results have practical significance for the automation of complex system management across various fields, including engineering, economics, and medicine**.**

BIBLIOGRAPHY:

1. Jin X., Zhang J., Su T., Bai Y., Kong J., Wang X. Modeling and analysis of data-driven systems through computational neuroscience wavelet-deep optimized model for nonlinear multicomponent data forecasting. *Computational Intelligence and Neuroscience*. 2021. URL: https://doi.org/10.1155/2021/8810046

2. Симонов Д. І., Заіка Б. Ю. Моделювання управління складними інформаційними багатокомпонентними системами. *Науковий вісник Ужгородського університету. Серія «Математика і інформатика»*. 2024. Вип. 44(1). С. 168–174. URL: https://doi.org/10.24144/2616-7700.2024.44(1).168-174

3. Daraghmeh M., Agarwal A., Jararweh Y. Optimizing serverless computing: A comparative analysis of multi-output regression models for predictive function invocations. *Simulation Modelling Practice and Theory*. 2024. Vol. 134. Article 102925. URL: https://doi.org/10.1016/j.simpat.2024.102925

4. Emami S. S., Martínez-Muñoz G. Deep learning for multi-output regression using gradient boosting. *IEEE Access*. 2024. Vol. 12. P. 17760–17772. URL: https://doi.org/10.1109/ACCESS.2024.3359115

5. Salehi F., Abbasi E., Hassibi B. The impact of regularization on high-dimensional logistic regression. *ArXiv*. 2019. Vol. abs/1906.03761. URL: https://doi.org/10.48550/arXiv.1906.03761

6. Tan C., Chen S., Ji G., Geng X. Multilabel distribution learning based on multioutput regression and manifold learning. *IEEE Transactions on Cybernetics*. 2020. Vol. 52. P. 5064–5078. URL: https://doi.org/10.1109/TCYB.2020.3026576

7. Khodarahmi M., Maihami V. A review on Kalman filter models. *Archives of Computational Methods in Engineering*. 2022. Vol. 30. P. 727–747. URL: https://doi.org/10.1007/s11831-022-09815-7

8. Guo L., Chen W., Liao Y., Liao H., Li J. Y. An edge-preserved image denoising algorithm based on local adaptive regularization. *Journal of Sensors*. 2016. Vol. 2016. Article ID 2019569:1–2019569:6. URL: https://doi.org/10.1155/2016/2019569

9. Tanaka G., Matsumori T., Yoshida H., Aihara K. Reservoir computing with diverse timescales for prediction of multiscale dynamics. *Physical Review Research*. 2022. Vol. 4, Iss. 3. Article L032014. URL: https://doi.org/10.1103/PhysRevResearch.4.L032014

10. Петрик Б. В., Неласа Г., Дубровін В. Аналіз часових послідовних потоків даних мережевого трафіку на основі вейвлет-перетворення. *Прикладні питання математичного моделювання*. 2020. Т. 3, № 1. С. 168–177. URL: https://doi.org/10.32782/2618-0340/2020.1-3.17

11. Salih A. M., Raisi-Estabragh Z., Galazzo I. B., Radeva P., Petersen S. E., Lekadir K., Menegaz G. A perspective on explainable artificial intelligence methods: SHAP and LIME. *Advanced Intelligent Systems*. 2024. Article 2400304. URL: https://doi.org/10.1002/aisy.202400304

12. Wu Y., Zhou Y. Prediction and feature analysis of punching shear strength of two-way reinforced concrete slabs using optimized machine learning algorithm and Shapley additive explanations. *Mechanics of Advanced Materials and Structures*. 2023. Vol. 30. P. 3086–3096. URL: https://doi.org/10.1080/15376494.2022.2068209

13. Shin J. Feasibility of local interpretable model-agnostic explanations (LIME) algorithm as an effective and interpretable feature selection method: comparative

fNIRS study. *Biomedical Engineering Letters*. 2023. Vol. 13. P. 689–703. URL: https://doi.org/10.1007/s13534-023-00291-x

14. Symonov D., Symonov Y. Methods for selecting models of functioning of multicomponent information and environmental systems. *Scientific Journal «Mathematical Modeling»*. 2024. Vol. 1, No 50. P. 57–63. URL: https://doi.org/10.31319/2519-8106.1(50)2024.304943

15. Симонов Д. І. Метод ентропії як інструмент оптимізації складних систем. *Журнал обчислювальної та прикладної математики*. 2024. № 1. С. 49–58. URL: https://doi.org/10.17721/2706-9699.2024.1.04

16. Do N., Hoang V., Doan V. A novel non-profiled side channel attack based on multi-output regression neural network. *Journal of Cryptographic Engineering*. 2023. Vol. 14. P. 427–439. URL: https://doi.org/10.1007/s13389-023-00314-4

17. Li C., Rakitsch B., Zimmer C. Safe active learning for multi-output Gaussian processes. *ArXiv*. 2022. Vol. abs/2203.14849. URL: https://doi.org/10.48550/arXiv.2203.14849

18. Park M., Choi Y., Lee N., Kim D. SpReME: Sparse regression for multi-environment dynamic systems. *ArXiv*. 2023. Vol. abs/2302.05942. URL: https://doi.org/10.48550/arXiv.2302.05942

19. Choe B., Kang T., Jung K. Recommendation system with hierarchical recurrent neural network for long-term time series. *IEEE Access*. 2021. Vol. 9. P. 72033–72039. URL: https://doi.org/10.1109/ACCESS.2021.3079922

20. Zhou X., Zhai N., Li S., Shi H. Time series prediction method of industrial process with limited data based on transfer learning. *IEEE Transactions on Industrial Informatics*. 2023. Vol. 19. P. 6872–6882. URL: https://doi.org/10.1109/TII.2022.3191980

21. Wang C., Chen X., Wu C., Wang H. AutoTS: Automatic time series forecasting model design based on two-stage pruning. *ArXiv*. 2022. Vol. abs/2203.14169. URL: https://doi.org/10.48550/arXiv.2203.14169

22. Tsanas A., Xifara A. Energy efficiency [Dataset]. *UCI Machine Learning Repository*. 2012. URL: https://doi.org/10.24432/C51307

REFERENCES:

1. Jin, X., Zhang, J., Su, T., Bai, Y., Kong, J., & Wang, X. (2021). Modeling and Analysis of Data-Driven Systems through Computational Neuroscience Wavelet-Deep Optimized Model for Nonlinear Multicomponent Data Forecasting. *Computational Intelligence and Neuroscience, 2021*. URL: https://doi.org/10.1155/2021/8810046

2. Symonov, D. I., & Zaika, B. Y. (2024) Modelling the management of complex information multicomponent systems. Scientific Bulletin of Uzhhorod University. Series 'Mathematics and Informatics', 44(1), 168-174. URL: https://doi.org/10.24144/2616-7700.2024.44(1).168-174 [in Ukrainian].

3. Daraghmeh, M., Agarwal, A., & Jararweh, Y. (2024). Optimizing serverless computing: A comparative analysis of multi-output regression models for predictive function invocations. *Simul. Model. Pract. Theory, 134*, 102925. URL: https://doi.org/10.1016/j.simpat.2024.102925

4. Emami, S.S., & Martínez-Muñoz, G. (2024). Deep Learning for Multi-Output Regression Using Gradient Boosting. IEEE Access, 12, 17760-17772. URL: https://doi.org/10.1109/ACCESS.2024.3359115

5. Salehi, F., Abbasi, E., & Hassibi, B. (2019). The Impact of Regularization on High-dimensional Logistic Regression. ArXiv, abs/1906.03761. URL: https://doi.org/10.48550/arXiv.1906.03761

6. Tan, C., Chen, S., Ji, G., & Geng, X. (2020). Multilabel Distribution Learning Based on Multioutput Regression and Manifold Learning. *IEEE Transactions on Cybernetics, 52*, 5064-5078. URL: https://doi.org/10.1109/TCYB.2020.3026576

7. Khodarahmi, M., & Maihami, V. (2022). A Review on Kalman Filter Models. *Archives of Computational Methods in Engineering, 30*, 727-747. URL: https://doi.org/10.1007/s11831-022-09815-7

8. Guo, L., Chen, W., Liao, Y., Liao, H., & Li, J.Y. (2016). An Edge-Preserved Image Denoising Algorithm Based on Local Adaptive Regularization. *J. Sensors, 2016*, 2019569:1-2019569:6. URL: https://doi.org/10.1155/2016/2019569

9. Tanaka, G., Matsumori, T., Yoshida, H., & Aihara, K. (2022). Reservoir computing with diverse timescales for prediction of multiscale dynamics. *Physical Review Research*, *4*(3), L032014. URL: https://doi.org/10.1103/PhysRevResearch.4.L032014

10. Petryk, B. V., Nelasa, G., & Dubrovin, V. (2020). Analysis of time sequential network traffic data streams based on wavelet transform. Applied issues of mathematical modelling, 3(1), 168-177. URL: https://doi.org/10.32782/2618-0340/2020.1-3.17 [in Ukrainian].

11. Salih, A. M., Raisi-Estabragh, Z., Galazzo, I. B., Radeva, P., Petersen, S. E., Lekadir, K., & Menegaz, G. (2024). A perspective on explainable artificial intelligence methods: Shap and lime. *Advanced Intelligent Systems*, 2400304. URL: https://doi.org/10.1002/aisy.202400304

12. Wu, Y., & Zhou, Y. (2023). Prediction and feature analysis of punching shear strength of two-way reinforced concrete slabs using optimized machine learning algorithm and Shapley additive explanations. *Mechanics of Advanced Materials and Structures, 30*, 3086 – 3096. URL: https://doi.org/10.1080/15376494.2022.2068209

13. Shin, J. (2023). Feasibility of local interpretable model-agnostic explanations (LIME) algorithm as an effective and interpretable feature selection method: comparative fNIRS study. *Biomedical Engineering Letters, 13*, 689 – 703. URL: https://doi.org/10.1007/s13534-023-00291-x

14. Symonov, D., Symonov, Y. (2024). Methods for selecting models of functioning of multicomponent information and environmental systems. Scientific Journal «Mathematical Modeling», Vol. 1, No 50, P. 57-63. URL: https://doi.org/10.31319/2519-8106.1(50)2024.304943

15. Symonov D.I. (2024). Entropy method as a tool for optimising complex systems. Journal of Computational and Applied Mathematics, No. 1, pp. 49-58. URL: https://doi.org/10.17721/2706-9699.2024.1.04 [in Ukrainian]

16. Do, N., Hoang, V., & Doan, V. (2023). A novel non-profiled side channel attack based on multi-output regression neural network. *J. Cryptogr. Eng., 14*, 427-439. URL: https://doi.org/10.1007/s13389-023-00314-4

17. Li, C., Rakitsch, B., & Zimmer, C. (2022). Safe Active Learning for Multi-Output Gaussian Processes. *ArXiv, abs/2203.14849*. URL: https://doi.org/10.48550/arXiv.2203.14849

18. Park, M., Choi, Y., Lee, N., & Kim, D. (2023). SpReME: Sparse Regression for Multi-Environment Dynamic Systems. *ArXiv, abs/2302.05942*. URL: https://doi.org/10.48550/arXiv.2302.05942

19. Choe, B., Kang, T., & Jung, K. (2021). Recommendation System With Hierarchical Recurrent Neural Network for Long-Term Time Series. *IEEE Access, 9*, 72033-72039. URL: https://doi.org/10.1109/ACCESS.2021.3079922

20. Zhou, X., Zhai, N., Li, S., & Shi, H. (2023). Time Series Prediction Method of Industrial Process With Limited Data Based on Transfer Learning. *IEEE Transactions on Industrial Informatics, 19*, 6872-6882. URL: https://doi.org/10.1109/TII.2022.3191980

21. Wang, C., Chen, X., Wu, C., & Wang, H. (2022). AutoTS: Automatic Time Series Forecasting Model Design Based on Two-Stage Pruning. *ArXiv, abs/2203.14169*. URL: https://doi.org/10.48550/arXiv.2203.14169

22. Tsanas, A. & Xifara, A. (2012). Energy Efficiency [Dataset]. UCI Machine Learning Repository. URL: https://doi.org/10.24432/C51307.