Research on modeling and algorithm of resource allocation optimization based on multi-factor constraints

Gonghao Li *

College of Information and Control Engineering, Qingdao University of Technology, Qingdao, China, 266520

* Corresponding Author Email: 18561161862@163.com

Abstract. As resource constraints and environmental pressures intensify, optimising systems under these constraints has become important in many fields, e.g. agriculture, energy scheduling and manufacturing. This requires balancing land, water, energy and so on. While existing optimisation methods address some single-constraint problems, they are poor at responding to dynamic changes in complex, multi-constraint systems due to poor adaptability and efficiency.In this paper, we propose a forecasting-scheduling modelling approach that identifies influencing factors using regression and growth models, and simulates system responses to resource constraints. It then combines mathematical scheduling methods for optimal resource allocation.Experimental findings show this framework can efficiently optimise resource allocation under these constraints, especially in constrained environments, with strong adaptability and stability.The growth model effectively reflects non-linear behaviour and the scheduling model significantly improves system efficiency, while respecting constraints.This study provides scientific decision support for scheduling resources, and has potential for industry-wide applications, and for real industrial and social systems.

Keywords: Data-driven, Optimal Resource Scheduling, Linear Regression, Logistic-Growth Models, Multi-Objective Planning.

1. Introduction

Modern engineering systems face increasingly complex challenges in scheduling resources, especially with multiple objectives and dynamic changes. The efficient allocation of resources (time, energy, space, money) is crucial for balancing output and cost, efficiency and environment, but is a pressing issue. Traditional optimisation methods, such as regression, linear programming and heuristic algorithms [1]-[17], have limitations like insufficient accuracy and lack of global optimaity in complex systems and dynamic environments. They also fail to effectively handle multi-objective optimization and dynamic changes. Conventional scheduling strategies are often sub-optimal, leading to inefficient resource allocation. These methods are inadequate for handling engineering problems with multiple objectives and dynamic constraints.

This paper proposes an integrated optimisation model to handle multi-constraint systems. Firstly, we identify relationships between input variables and response indicators (linear and non-linear), then simulate system behaviour under multiple constraints using an accurate prediction model. Next, we design an integrated optimisation strategy combining mathematical planning methods to ensure resource-constrained system benefits are maximised. Our dynamic resource allocation strategy resolves input perturbations, effectively dealing with system conflicts and improving resource utilisation.

The method has a wide range of application prospects, especially for production line scheduling, energy management, supply chain optimisation and intelligent manufacturing. The model is general and scalable, supporting practical engineering management and intelligent decision-making systems. Technological progress will improve the framework's support for optimising and managing complex systems, solving problems with limited resources, conflicting objectives and dynamic perturbations.

2. Related work

In complex system modelling, identifying causality and predicting behaviour based on historical data is a key research area. Traditional linear regression models are widely used due to their simplicity and interpretability, especially in simple systems. However, they struggle to accurately capture complex change trends in nonlinear dynamic systems. Researchers have therefore proposed nonlinear models, such as logistic growth models, S-functions and polynomial regression, to improve adaptability to saturation effects and constrained growth. For example, a stochastic SIRS infectious disease transmission model was constructed based on the logistic model to portray the disease evolution process; a prediction of China's demographic evolution trend was fitted; and multivariate nonlinear regression was used to investigate product quality fluctuation characteristics under different raw material conditions [18]-[20]. These methods have achieved some success, but face challenges when dealing with complex features.

There are many issues around how we use and schedule resources in engineering. Constraints based on capacity, emission, and rationing have traditionally guided these. Methods like Mixed Integer Linear Programming (MILP) solve complex decision problems, but are limited by real-time changes and multi-stage scheduling problems, especially for high-dimensional combinatorial optimisation, which requires heuristic approaches.

Machine learning and AI have led to a growing focus on 'prediction-driven optimisation' in resource scheduling. This approach merges machine learning with traditional optimisation methods. Instead of treating prediction and scheduling as separate processes, this approach integrates them, which excels in dynamic environments. However, research is focused on specific aspects. This paper proposes a data-driven multi-stage modelling scheme that links identification, prediction, and scheduling through a unified framework, enhancing the system's adaptability and efficiency in complex environments.

3. Research Methodology

3.1. Overall framework

This paper presents a multi-stage algorithm for optimising the use of resources in constrained scenarios (Fig. 1). The process is data-driven, starting with identifying key factors from inputs and establishing mapping relationships. It then models system behaviour as a nonlinear growth model and uses this analysis to predict outcomes. Lastly, mathematical planning methods optimise the resource allocation to maximise benefits within given limits. The framework integrates analysis, modelling and resource scheduling.

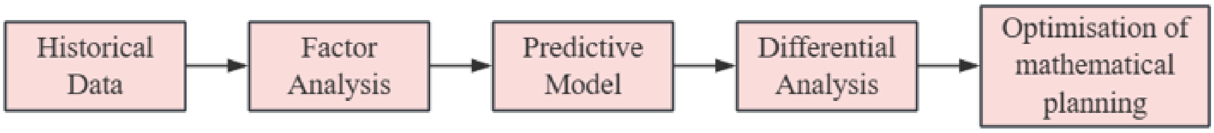

Figure 1. Algorithm flow chart of the system.

3.2. Input factor modelling (Variable selection + Regression analysis)

3.2.1. Methodological steps

The key factors affecting the system output are identified using a multidimensional input model and prediction model. Data is initially segmented using cluster analysis to identify potential factors. These factors are then identified using statistical methods such as Pearson's coefficient and analysis of variance (ANOVA). A linear regression model is used to establish the mapping relationship between inputs and outputs. During this stage, special attention must be paid to variance inflation factor (VIF) to ensure the model's stability and reliability. The goodness of fit of the model is tested to ensure its validity and predictive accuracy.

3.2.2. Analysis of K-means clustering theory

The K-Means clustering algorithm [21] is a popular unsupervised learning algorithm used in data analysis and pattern recognition. The goal is to divide the dataset into K clusters so that samples within each cluster are similar and different clusters are dissimilar.

The K-Means algorithm adjusts the centres of the clusters by means of iteration to obtain a suitable clustering result.

Initial centres can be randomly selected or heuristics can be used. Data points are then assigned to the nearest cluster centre based on the current centre. The distances are usually calculated using the Euclidean distance.

$$d(x,c_k)=\sqrt{\sum_{i=1}^{n}(x_i-c_{k,i})^2} \tag{1}$$

Where x is the data point, c_k is the cluster centre and n is the dimension of the data.

The cluster centres are recalculated based on the mean value of all data points in each cluster:

$$c_k=\frac{1}{|S_k|}\sum_{x\in S_k}x \tag{2}$$

Where $|S_k|$ the number of data points in cluster k is, S_k is the set of data points in cluster k, and x is a data point. Repeat the above steps until the centre of the cluster does not change anymore or a predefined number of iterations is reached.

3.2.3. Linear regression models and related theory

Linear regression [22] is a regression analysis method used in statistics, machine learning and other fields. It describes the linear relationship between the independent variable (input) and the dependent variable (output) by fitting a straight line. A simple linear regression model assumes a linear relationship between the dependent and independent variables, i.e., the dependent variable can be predicted by the weighted sum of the independent variables.

$$y=\beta_0+\beta_1 x+\varepsilon \tag{2}$$

y is the dependent variable (output) and x is the independent variable (input).

β_0 Is the intercept term (bias), which indicates the value of y when x=0.

β_1 Is the regression coefficient (slope), which indicates the degree of influence of the independent variable x on the change in the dependent variable y?

ε Is the error term (noise), indicating other influences that the model fails to capture.

3.3. Modelling of system growth projections (non-linear modelling)

3.3.1. Methodological steps

In the second stage, the system's nonlinear response characteristics begin to emerge. At resource input onset, the system typically exhibits rapid growth, which saturates at a certain limit, resulting in an S-shaped curve. Many engineering systems exhibit this behaviour, therefore, we employ a logistic model to simulate the system's growth process under resource constraints. We first analysed the residuals of the linear regression model to gauge its effectiveness and found that logistic models were a more suitable tool. By fitting the logistic model parameters via least squares and differential analysis, we can accurately calculate the system's maximum rate of growth and its corresponding resource inputs, enabling us to determine the system's maximum response and resource scheduling.

3.3.2. Logistic Model Theoretical Analysis

The logistic growth model [23]-[24] is a nonlinear model describing the growth process of populations and resources in a limited environment. It is used to describe real-world growth processes where saturation limits exist. Unlike the exponential growth model, the logistic model considers resource limitations and gradually slows growth rate until it reaches a maximum value (saturation).

$$P(t) = \frac{K}{1 + \left(\frac{K - P_0}{P_0}\right) e^{-rt}} \tag{4}$$

Parameter description:

K: maximum system carrying capacity. r: growth rate. t : resource input variable or time variable.

The methodological framework in this paper provides comprehensive models and optimal solutions for multi-faktor/multi-constraint systems through a three-phase strategy of 'data + non-linear prediction + optimal scheduling', applicable to all kinds of engineering resource management scenarios.

4. Results

4.1. Purpose and Setting of the Experiment

In this paper, we validate the algorithm through the 'Agricultural Resource Scheduling' scenario to show the applicability of the integrated prediction-optimisation modelling method. Historical region data 2000-2022 is used for the experiments, which include system inputs like average annual temperature, sunshine hours, and precipitation. Despite the data sources being from the agricultural sector, the central objective is to demonstrate the generalisability of the methodology. The experimental results show that the model can predict the relationship between input factors and response variables, and optimally allocate resources. This validates the method's application in agricultural scenarios, and demonstrates its wide applicability.

4.2. Linear regression modelling and analysis of variance

Variable selection and cluster analysis were used to identify key input factors. A cluster analysis of the data extracted three candidates: average temperature, sunshine hours, and annual precipitation. These were correlated with the response variable, Y (system growth rate), measured using different conditions. SPSSAU software was used for clustering and three groups were taken.

The samples were classified using K-means cluster analysis. The final clustering resulted in three groups, accounting for 40.91%, 22.73% and 36.36% of the total. The distribution of the three classes is more uniform, indicating a better clustering effect.

A linear regression model (e.g., Equation 8) was established for describing the relationship between the input factors and the response variables. The established model is shown in Figure 2.

$$Y = \beta_0 X_1 + \beta_1 X_2 + \beta_2 X_3 + \varepsilon \tag{5}$$

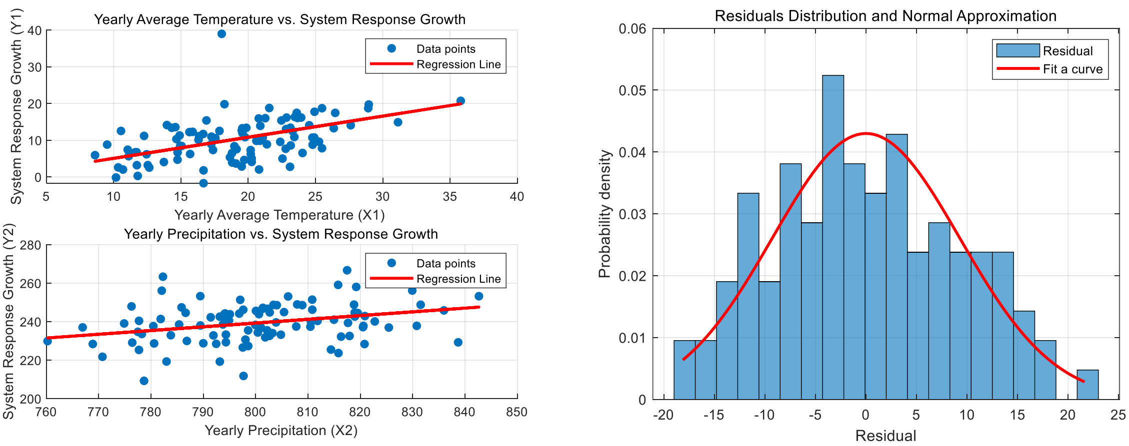

Figure 2. Linear regression model diagram and Plot of residual analysis results.

The model and data are imported into SPSS software, which performs linear regression analysis to explore the influence of each factor on the system response and the relationship between the data. The regression analysis module provides detailed statistical indicators such as regression coefficients, goodness-of-fit, significance level, etc., which help to assess the validity and predictive ability of the model. SPSS improves the calculation speed and ensures the reliability and accuracy of the analysis results.

A scatter plot of annual grain growth rate vs. annual precipitation shows linear relationship, as does the same for annual grain growth rate and precipitation. Scatter plots for annual grain growth rate vs. average annual temperature also show a linear relationship.

The results show an R^2 of 0.65, indicating that the model explains 65% of the variation in the response. All VIF values are less than 2, indicating no serious covariance problem.

DW value of 2.286 is close to 2, indicating a well-constructed model without an autocorrelation problem.

Normal means that the residuals obey a normal distribution, and its variance $\sigma^2 = var\ (ei)$ reflects the accuracy of the regression model, generally the smaller the σ, the higher the accuracy of the obtained regression model to predict y. According to the results Figure 2 residuals obey normal distribution indicates that the assumptions of the regression model are reasonable, the estimation of the regression coefficients is unbiased and optimal, and it is possible to carry out effective hypothesis testing and confidence interval calculation.

4.3. Logistic growth model fit

We used a Logistic growth model to fit the 'S'-shaped growth behaviour of the system's response to gradually increasing input, simulating the effect of resource input on that response. The aim of the experiment is to demonstrate the full modelling process of the system response from rapid growth to smooth saturation. By fitting historical data, we captured slowing growth and predicted the system's saturated state under resource constraints. The established Logistic growth model is shown in Fig. 3.

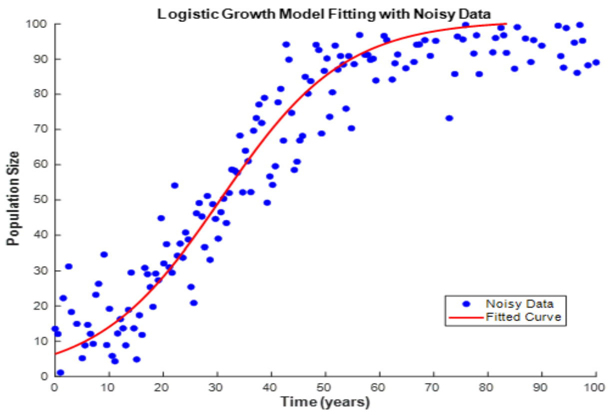

Figure 3. Logistic growth model diagram.

The model fits the trend of slowing growth in historical data well, especially when resources reach maximum capacity. Predicted values are less than 5% different from true values, showing the model is accurate and captures the system process from rapid growth to smooth saturation. This verifies the Logistic growth model's effectiveness in describing system response and supports optimal scheduling.

4.4. Optimal scheduling solution results

A mathematical planning model was used to analyse how to grow food in the region. The aim was to maximise food production and cultivation. The model considered various constraints and objectives.

We choose a representative region, and the following are the water demand per hectare, food production per hectare, GHG emissions per hectare, and the corresponding food price for food crops in that region, respectively:

The above model was entered into LINGO to value the custom parameters and the optimal results obtained are shown in Table 1.

Table 1. Optimal scheduling solution result table.

Types	Wheat (ha)	Rice (ha)	Maize (ha)	Maximum grain production (tonnes)	Maximum amount (yuan)
Plantation	5.00	0.00	4.50	47.87	9985.6
	6.73	0.00	3.27	44.63	11703.3

The results show that rice is uncompetitive due to its high water and greenhouse gas use. Corn is competitive in terms of grain yield, but less so in terms of profit, due to its lower price. Wheat is less competitive than maize, but more so than rice, due to the lower ratio of staple crops planted. Wheat is more competitive than maize due to its higher price. This experiment uses agricultural data to verify the modelling and scheduling method proposed in this paper. The model shows good prediction accuracy and optimisation ability under multiple constraints, indicating that it can be widely extended to manufacturing, energy, logistics and other systems as a general resource allocation decision tool.

5. Conclusions

In this paper, we present a method for solving the optimal scheduling of limited resource systems. We then verify the method's adaptability and optimisation capability through experiments. These show that the model can effectively achieve optimal resource allocation in environments with limited resources. The model also has potential for a wide range of cross-industry applications. With technological advances, the dynamic system modelling approach will be combined with the scheduling process. This will enhance the system's optimisation capability in complex environments. Combining intelligent optimisation algorithms will extend the application scope of the model in complex systems. It will also enhance the ability of global optimal solutions.

Despite the good results of this research, it is still facing certain limitations and challenges. Firstly, it is more dependent on the quality of the input data. In future, improvements are needed to the method for dealing with missing or unreliable data and with data that is subject to change over time. To deal with these challenges, we may introduce time series models (e.g. LSTM) or robust optimisation methods (e.g. robust optimisation) in future. It is also difficult to include environmental factors when there are high-dimensional inputs, so further research is required to improve their quantification and modelling. This study provides an effective framework for predictive modelling and optimising resources that can be easily scaled up for systems that are subject to multiple constraints. As technology develops and better algorithms are created, the framework will be optimised to enhance its dynamic modelling capability and intelligence level and be able to cope with increasingly complex systems. It will also be able to provide support for optimisation in a wider range of engineering and social systems.

References

[1] Liang Shuhong, Chen Xueke, Kang Xiaoyi. Research on the Influencing Factors of University Empowering Secondary Faculties - An Empirical Analysis Based on the Influencing Factors of University Faculty Resource Allocation [J]. Journal of Higher Education, 2025, 11(10):1-8.

[2] Qi RuiXin. Research on optimisation of foreign trade logistics network of agricultural products based on heuristic algorithm [J]. National Circulation Economy, 2025, (06):12-15.

[3] LU Natural, LIU Jiang, NI Feng, et al. Research on the impact of corporate governance structure on ESG responsibility fulfilment [J/OL]. Operation and Management, 1-16[2025-04-19].

[4] Wu Huimin. Research on cost overrun factors of engineering management projects based on regression analysis [J]. Heilongjiang Science, 2025, 16(04):147-149.

[5] ZHANG Shuai, ZHANG Guojiao, HUANG Jiahua. Research on the impact of social capital on the community integration of returning entrepreneurial college students--Based on binary logistic regression analysis [J]. SAR Economy, 2025, (01):82-85.

[6] Wang Kun. Application of linear programming in surface coal mine transport planning [J]. Mining Equipment, 2025, (01):68-70.

[7] CHEN Xueshuo, MAO Yuxing, XU Yihang, et al. Task offloading method for smart terminals in resource-constrained environments [J/OL]. Computer Engineering and Application, 1-10[2025-04-19].

[8] LUO Liang, LUO Kaiyuan, ZHANG Zhehao, et al. Application of linear programming simplex method in mine ore allocation [J]. Xinjiang Iron and Steel, 2024, (04):14-16.

[9] ZHAO Zi-Long, WU Peng, SUN Shao-Long, et al. Optimisation of emergency rescue scheduling under major forest fires considering limited resources[J/OL]. Systems Engineering Theory and Practice, 1-21[2025-04-19].

[10] Zeng Wenlong, Peng Fei, Feng Huan, et al. Resource optimisation model of power engineering construction system and its application [J]. Construction Economy, 2024, 45(S1):442-447.

[11] XU Xiuyuan, ZHOU Hao. Research on land use optimization in the Pearl River Delta region based on grey prediction and linear programming [J]. Land Resources Guide, 2024, 21(02):79-85.

[12] YANG Xiaolong, XU Kun, YI Jianbo, et al. Frequency planning of shortwave diversity communication network based on heuristic algorithm. Communication Technology, 2025, 58(01):63-67.

[13] WU Liangwen, ZHEN Lu. Research on optimisation model and algorithm of major engineering construction management based on drones [J]. Frontiers of Engineering Management Science and Technology, 2024, 43(06):58-67.

[14] ZHAO Lilin, ZHOU Ying, YANG Shuyan. Influencing factors of fear of disease progression in elderly hypertensive cerebral haemorrhage patients based on multiple linear regression analysis [J]. General Practice Nursing, 2025, 23(07):1330-1334.

[15] ZHAO Chunmei, ZHU Xiong, FENG Yunyu. Diagnostic study on linear regression analysis of coal seam gas content and pressure with burial depth [J]. Coal Technology, 2025, 44(04):218-221.

[16] FENG Daguang, FENG Sizhe, LI Zhaoxing, et al. Comparison of temperature prediction methods for solar greenhouse based on multiple linear regression model[J/OL]. Journal of Shenyang Normal University (Natural Science Edition), 1-7[2025-04-20].

[17] HAN Shuo, WANG Bin, LIU Sha. Research on predicting equivalent spherical lenses for myopic children in primary and secondary schools based on multiple linear regression analysis [J]. Journal of Hubei Institute of Science and Technology (Medical Edition), 2025, 39(02):147-151.

[18] ZHAO Yanjun, SUN Xiaohui, SU Li, et al. A stochastic SIRS infectious disease model with logistic growth and containing Ornstein-Uhlenbeck process and psychological effects [J]. Journal of Northeast Normal University (Natural Science Edition), 2024, 56(04):1-10.

[19] YANG Shuang, CHEN Hongwu, LIU Mingyang, et al. Study on population forecasting based on improved logistic blocked growth model [J]. Inner Mongolia Science and Economy, 2024, (02):29-33+50.

[20] TANG Pengxiang, YANG Neng, ZHANG Xiaomei. Ore processing quality prediction and control based on multivariate nonlinear regression and BP neural network [J]. Microcomputer Applications, 2025, 41(01):43-48.

[21] WANG Bingshen, WANG Guochang, WEI Yanhua. Improvement of the method for determining the number of clusters based on k-means algorithm [J/OL]. Statistics and Decision Making, 2025, (07):59-64[2025-04-19].

[22] FENG Daguang, FENG Sizhe, LI Zhaoxing, et al. Comparison of solar greenhouse temperature prediction methods based on multiple linear regression models [J/OL]. Journal of Shenyang Normal University (Natural Science Edition), 1-7[2025-04-19].

[23] LU Si-Yu, KUANG Zhao-Min, LI Li. Growth and development of sugarcane plant height in Guangxi based on logistic model [J/OL]. Tropical Agricultural Science, 1-6[2025-04-19].

[24] XIAO Chun'an, CAI Jiabing, CHANG Hongfang, et al. Precision screening and differentiation of farmland surface temperature data considering crop growth status [J]. Journal of Agricultural Engineering, 2022, 38(22):89-101.