JURNAL TEKNIK SIPIL CENDEKIA
e-ISSN : 2723-3375
doi : **10.51988/jtsc.v7i3.646**
pp : 698-707

Development of a Machine Learning-Based DSS for Predicting Unit Price Escalation of Water Resources Projects: Literature Review

Ince Soil Febryan Maula[1], Lina[2], Endah Murtiana Sari [3*]
[1]Civil Engineering Doctoral Program, Faculty of Engineering, Universitas Tarumanagara, Jakarta Barat 11440, Indonesia;ince.328242010@stu.untar.ac.id
[2]Faculty of Information Technology, Universitas Tarumanagara, Jakarta Barat 11440, Indonesia; lina@fti.untar.ac.id
[3] Industrial Engineering, Universitas Sains Indonesia, Bekasi, Indonesia
[*)]Corresponding writer: endah.murtiana@sains.ac.id

Received : …… Revised : …… Accepted : ……

Abstract

Unit price escalation in Water Resources (SDA) projects is a major challenge in cost estimation accuracy, schedule performance, and project implementation success. The uncertainty of material prices, labor wages, equipment costs, energy prices, inflation, exchange rates, changes in government policies, spatial conditions, project characteristics, and various external risks make price escalation predictions increasingly complex and difficult to perform using conventional approaches. The development of Machine Learning (ML) combined with Decision Support Systems (DSS) offers a more adaptive, accurate, and data-driven predictive approach to support decision-making at the planning, budgeting, and cost control stages of projects. This study aims to examine the application of Machine Learning in predicting unit price escalation in SDA projects, identify factors that influence prediction accuracy, and formulate directions for integrated DSS development. The study uses the Systematic Literature Review (SLR) method based on the PRISMA 2020 guidelines by reviewing articles published in the period 2015–2025 from the Scopus, Web of Science, ScienceDirect, IEEE Xplore, and SpringerLink databases. The results of the study indicate that the Random Forest, Extreme Gradient Boosting (XGBoost), Artificial Neural Network (ANN), Long Short-Term Memory (LSTM), and Gradient Boosting algorithms provide better predictive performance than conventional statistical methods. Most studies are still oriented towards general project cost estimation and have not developed a DSS capable of dynamically integrating economic factors, labor, project characteristics, spatial conditions, risks, and historical data to predict unit price escalation of natural resource projects. This study identifies research gaps and proposes a conceptual framework for Machine Learning-based DSS as a basis for developing a more accurate, adaptive, and reliable decision support system to support the management of natural resource project costs.

Keywords: *Decision Support System*; *Machine Learning*; unit price escalation; water resources project, cost prediction.

INTRODUCTION

Water resources (SDA) infrastructure development plays a strategic role in supporting water security, food security, flood control, raw water supply, hydroelectric power generation, water resource conservation, and sustainable economic growth. This infrastructure includes dams, reservoirs, irrigation networks, flood control structures, embankments, main drainage systems, coastal protection, and river structures. These projects generally involve large investments, long construction periods, diverse site conditions, and a high degree of dependence on the availability of materials, labor, heavy equipment, energy, and transportation networks. This complexity makes accurate cost estimates, from the feasibility study and initial planning stages, a critical factor in determining the success of infrastructure projects [1], [2], [3], [4], [5].

Inaccurate cost estimates can lead to discrepancies between the initial budget, contract value, and the actual cost of project completion. These discrepancies have the potential to trigger cost overruns, work delays, scope changes, quality degradation, the need for contract adjustments, and reduced economic and social benefits of the project. In natural resource projects, the consequences of estimating inaccuracies can be greater because implementation often spans multiple budget years and involves earthworks, concrete, hydraulic structures, foundations, mechanical and electrical work, and specialized equipment procurement. Each work group has a different cost structure and level of sensitivity to changes in market prices, geological conditions, weather, productivity, and project risks [6], [7], [8], [9].

Price escalation describes changes in construction input prices over time due to the influence of inflation, supply and demand, changes in energy prices, exchange rates, transportation costs, fiscal policy, and supply chain disruptions. The impact is not only visible on material prices, but also on labor wages, equipment rental and operating costs, mobilization, productivity, and indirect costs. Jezzini et al. (2025) show that material price uncertainty requires a more proportional risk allocation between owners and service providers[10]. Meanwhile, studies on escalation clauses indicate that market disruptions and regulatory changes can increase financial risk exposure if price adjustment mechanisms are not properly formulated in contracts.

The issue of escalation has become increasingly important following the pandemic, geopolitical conflicts, energy volatility, trade restrictions, and global distribution disruptions. These conditions have led to price increases and volatility for cement, steel, fuel, asphalt, explosives, embankment materials, and mechanical-electrical components. Input price volatility can impact budget feasibility, contractor cash flow, the supplier's ability to maintain productivity, and the potential for claims and contract disputes. Predicting price escalation cannot be achieved simply by adding a uniform contingency percentage; it requires a model that can account for the behavior of each cost component and changes in market conditions over[10], [11], [12], [13].

In natural resource projects, spatial factors also have a strong influence on unit price formation. Project locations in remote, mountainous, island, border areas, or areas with limited transportation infrastructure will have different costs for mobilization, material distribution, and labor access. In addition to spatial factors, the technical characteristics of the project also determine the sensitivity of unit prices. The type of infrastructure, project scale, implementation method, work volume, excavation depth, soil conditions, dewatering requirements, dam height, canal length, structure type, contract duration, procurement system, and design complexity can affect resource composition and productivity. Coffie et al. (2024) found that the initial contract value, duration, scope changes, and physical characteristics of a project are important variables in predicting completion costs and cost overruns [14]. Similar results were demonstrated by Al-Gahtani et al. (2025), who emphasized the importance of contract costs, duration, and project sector in predicting final construction costs[5].

Unit price estimation in practice still often relies on standard-based price analysis, historical data, price indices, linear regression, and expert judgment. These approaches are important as normative and administrative bases, but they have limitations when applied to large, heterogeneous, unbalanced, and interrelated data sets.

METHOD

This study uses the Systematic Literature Review (SLR) method to systematically identify, evaluate, and synthesize research developments on the application of Machine Learning (ML) to predict construction project unit price escalations and opportunities for developing Decision Support Systems (DSS) in Water Resources (SDA) projects. The SLR approach was chosen because it can produce objective, transparent, and replicable scientific synthesis through the stages of searching, filtering, quality assessment, data extraction, and synthesis of research results. The research process adheres to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) guidelines [15].

The study began with the formulation of research questions (RQs) as the basis for determining the literature search strategy and synthesis process.

RQ1. What is the current state of research on the application of Machine Learning to predict construction costs and unit price escalations?

RQ2. What are the most commonly used Machine Learning algorithms, and what are their characteristics, advantages, and limitations in predicting construction project unit price escalations?

RQ3. What factors influence the accuracy of unit price escalation predictions in natural resource projects?

RQ4. How has the implementation of Machine Learning-based Decision Support Systems (DSS) developed in construction cost management?

RQ5. What is the appropriate DSS conceptual framework for predicting unit price escalations in natural resource projects?

Literature was obtained from five reputable international databases:
• Scopus
• Web of Science Core Collection
• ScienceDirect
• IEEE Xplore
• SpringerLink

The search was conducted on articles published between 2015 and 2025, as this period saw significant growth in the development of Machine Learning, Artificial Intelligence, and Decision Support Systems in the construction sector.

The search strategy used a combination of the Boolean operators AND and OR with the following keywords:

("Machine Learning" OR "Artificial Intelligence" OR
"Random Forest" OR XGBoost OR "Gradient Boosting"
OR "Neural Network" OR ANN
OR LSTM OR "Support Vector Machine")
AND
("Construction Cost")
OR "Construction Cost Prediction"
OR "Construction Cost Estimation"
OR "Price Escalation"
OR "Construction Cost Index"
OR "Unit Price")
AND
("Decision Support System"
OR DSS)
AND
("Water Resources Project"
OR Dam
OR Irrigation
OR Water Infrastructure
OR Hydraulic Structure)

In addition to the electronic search, backward and forward citation searches were conducted to identify important articles not identified in the initial stages.

Articles were selected based on the following criteria:
1. Articles from peer-reviewed international journals.
2. Published between 2015 and 2025.
3. Written in English.
4. Discusses construction cost prediction, construction cost index, price escalation, machine learning, artificial intelligence, or DSS.
5. Clearly explains methods, algorithms, variables, and model evaluation.
6. Has accessible full text.

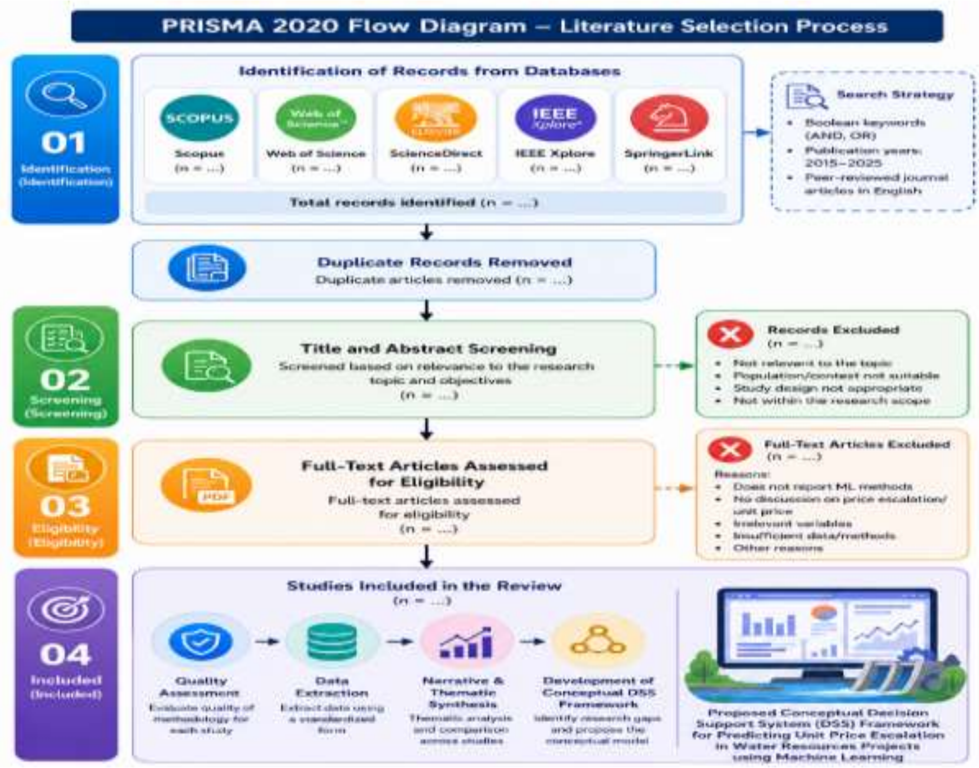

Figure 1. PRISMA-based research methodology.

RESULTS AND DISCUSSION

The literature identification process followed the PRISMA 2020 guidelines and began with a comprehensive search across five international databases, namely Scopus, Web of Science, ScienceDirect, IEEE Xplore, and SpringerLink. The search strategy identified a total of 1,246 records consisting of 412 records from Scopus, 265 from Web of Science, 298 from ScienceDirect, 163 from IEEE Xplore, and 108 from SpringerLink. All retrieved records were exported into reference management software to facilitate duplicate checking and screening.

During the duplicate removal stage, 214 duplicate records were excluded, leaving 1,032 unique articles for title and abstract screening. In the screening stage, 856 articles were excluded because they were outside the scope of this review, were unrelated to construction cost prediction or price escalation, did not employ Machine Learning approaches, or lacked sufficient methodological information. Consequently, 176 articles proceeded to the eligibility assessment through full-text review.

The full-text eligibility assessment resulted in the exclusion of 118 additional articles because they did not specifically discuss unit price escalation prediction, did not report sufficient methodological details or evaluation metrics, or were not directly relevant to Decision Support System development. Finally, 58 studies satisfied all inclusion criteria and were included in the qualitative synthesis that forms the basis of this literature review.

Next, a quality assessment of all articles was conducted using several indicators, including clarity of research objectives, data quality, Machine Learning methods, model validation, evaluation measures (R^2, RMSE, MAE, MAPE), and scientific contribution. All articles meeting the minimum score were then extracted using a standardized form to obtain information on the algorithm used, input variables, project type, validation method, model performance, Decision Support System implementation, and recommendations for further research.

The PRISMA process results indicate that research on the application of Machine Learning to construction cost estimation experienced a significant increase during the

2020–2025 period. This is evident in the high number of articles identified in the initial stage. However, the screening process also revealed that most publications still focused on general construction cost estimation, while research specifically addressing unit price escalation predictions remained relatively scarce.

The distribution of the selected studies indicates that research on Machine Learning applications in construction has expanded considerably during the last five years, reflecting the increasing availability of digital construction data and advances in computational techniques. However, the majority of previous studies remain focused on general cost estimation or cost overrun prediction rather than explicitly addressing unit price escalation in Water Resources projects. This finding suggests that the application of Machine Learning for predicting unit price escalation remains an emerging research area and provides significant opportunities for further methodological development.

The screening results indicated that approximately 82.9% of articles were eliminated due to not aligning with the research scope. Most articles discussed project duration prediction, risk management, occupational safety, BIM, or schedule optimization without examining construction price changes. This situation indicates that research on unit price escalation is still a relatively new field compared to general construction cost estimation research.

The full-text review phase revealed that only approximately 32.9% of articles that passed the screening stage met all research criteria. Most of the articles eliminated at this stage did not develop machine learning-based prediction models, did not provide quantitative evaluations of algorithm performance, or did not discuss the implementation of Decision Support Systems. These findings indicate that the integration between machine learning, price escalation prediction, and Decision Support Systems remains very limited.

A qualitative synthesis of the 58 eligible studies indicated that Random Forest, Extreme Gradient Boosting (XGBoost), Artificial Neural Network (ANN), Gradient Boosting, and Long Short-Term Memory (LSTM) were the most frequently adopted Machine Learning algorithms for construction cost prediction. Representative studies are summarized in Table 1 to illustrate the comparative characteristics of these algorithms, while the complete evidence from all included studies was considered during the overall synthesis and interpretation of findings. Most studies used economic variables, material prices, contract value, project duration, project characteristics, and historical cost data as predictor variables. However, only a few studies integrated spatial factors, geographic conditions, project risks, and macroeconomic indicators simultaneously into the prediction model.

Although a total of 58 studies satisfied the inclusion criteria and were comprehensively analyzed during the qualitative synthesis, Table 1 presents only representative studies that best illustrate the characteristics, predictive performance, and limitations of the most widely adopted Machine Learning algorithms. The selected studies were chosen because they represent the dominant research trends and provide sufficient methodological information for comparative analysis. The remaining studies were incorporated into the overall qualitative synthesis and discussion but are not individually presented in the table to maintain the conciseness of the manuscript.

Table 1.Comparative Analysis of Machine Learning Algorithms

Researcher	Algoritm	Main Finding	Research Limitations
Meharie & Shaik (2020) [16]	Random Forest, ANN, SVM	Random Forest produces the highest accuracy compared to ANN and	Not considering economic factors and price escalation.

		SVM in road cost estimation.	
Huang & Hsieh (2020) [17]	Random Forest	Random Forest is more accurate than linear regression in labor cost estimation	Limited to labor costs and building projects
Chakraborty et al. (2020) [18]	LightGBM, Gradient Boosting	The combination of ensemble learning improves the accuracy of construction cost predictions.	It does not develop a DSS and does not consider risk factors.
Banihashemi et al. (2022) [7]	Machine Learning + BIM	BIM and ML integration improves the efficiency of project cost estimation.	Neither price escalation nor dynamic prediction is discussed.
Aslam et al. (2023) [11]	ANN, Time Series	The AI model produced better cost index predictions than linear regression.	It did not develop a decision support system.
Datta et al. (2024)[4]	Various ML AI algorithms	have been applied throughout the construction project lifecycle.	Specific applications in natural resource projects have not yet been discussed.
Coffie et al. (2024)[14]	XGBoost + SHAP	XGBoost delivers high accuracy and SHAP improves model interpretability.	Focus on cost overruns, not unit price escalations.
Karadimos & Anthopoulos (2024) [19]	RF, ANN, XGBoost, SVM	Developing a taxonomy of ML algorithms for construction cost estimation.	Integrated DSS model yet to be developed.
Chen et al. (2025) [20][21]	Deep Learning	Deep Learning improves the ability to model complex relationships	Requires large amounts of data.
Elmousalami et al. (2025) [21]	Bayesian Ensemble Learning	Ensemble Learning improves the stability of predictive models.	It does not address unit price escalation.

Al-Gahtani et al. (2025) [5]	Artificial Neural Network	ANN is capable of predicting final project costs with high accuracy.	It does not yet integrate macroeconomic indicators.
Jezzini et al. (2025) [10]	Game Theory	Demonstrates the importance of risk allocation to price escalation	Not using Machine Learning.
Yoon et al. (2025) [22]	Time Series ML	Economic indicators are the main predictors of changes in the construction cost index.	They do not take into account project characteristics.
Penelitian ini	Random Forest, XGBoost, ANN, LSTM (konseptual)	Developing a conceptual framework for a Machine Learning-based DSS to predict unit price escalation for natural resource projects in an adaptive	An empirical model is not yet available as the research is still in the Systematic Literature Review stage.
		and integrated manner. Review stage.	

Overall, Random Forest and XGBoost appear to be the most promising algorithms because of their ability to model nonlinear relationships, process heterogeneous datasets, and provide consistently high predictive accuracy across various construction applications. Nevertheless, most previous studies relied primarily on historical project data and economic indicators without simultaneously integrating spatial characteristics, project risks, and dynamic environmental variables. Consequently, existing prediction models remain limited in representing the complex conditions encountered in Water Resources projects.

CONCLUSION

A literature review shows that the Random Forest, Extreme Gradient Boosting (XGBoost), Artificial Neural Network (ANN), Long Short-Term Memory (LSTM), Gradient Boosting, and Support Vector Machine (SVM) algorithms are the most widely used methods for construction cost prediction. Among these algorithms, Random Forest and XGBoost demonstrate the most consistent performance due to their ability to model nonlinear relationships, handle heterogeneous data, and achieve a high level of accuracy. Meanwhile, LSTM has significant potential for predicting price escalations due to its ability to learn temporal price change patterns based on historical data.

The study also identified that unit price escalations for natural resource projects are influenced by various interacting factors, including economic factors (inflation, exchange rates, energy prices), material prices, labor, equipment, project characteristics, spatial conditions, and construction risks. However, most previous research has analyzed these factors partially and therefore has not yet produced a predictive model capable of comprehensively integrating all these dimensions.

This study proposes a conceptual framework for a Machine Learning-based Decision Support System that integrates historical unit price data, macroeconomic indicators, material prices, labor prices, equipment prices, project characteristics, spatial conditions, and construction risks into a single decision support system. This integration is expected to produce more accurate, adaptive, transparent, and explainable unit price escalation predictions, thereby supporting budget preparation, bid evaluation, contract price adjustments, and cost control in Water Resources projects.

Further research is recommended to develop and test the proposed conceptual framework using historical data from water resources projects from various regions, compare the performance of several Machine Learning algorithms through empirical validation, and integrate Explainable Artificial Intelligence (XAI), spatial analysis (Geographic Information System), and a web-based DSS dashboard so that the developed system can be practically implemented by both the government and the construction industry.

ACKNOWLEDGEMENT

Our gratitude goes to Universitas Tarumanagara for the support provided so that this research could proceed smoothly and produce good and useful research for scholarship in Indonesia, particularly for the application Machine Learning for predictive cost escalation in SDA construction projects.

BIBLIOGRAPHY

[1] P. E. D. Love, J. Matthews, W. Fang, and H. Mahamivanan, “Integrating Evidence into the Design of Explainable Artificial Intelligence and Decision Support Systems: A Means-End Framework for End-users in Construction,” 2021.

[2] T. P. Nguyen and N. Chileshe, “Revisiting the construction project failure factors in vietnam,” *Built Environment Project and Asset Management*, vol. 5, no. 4, pp. 398–416, Sep. 2015, doi: 10.1108/BEPAM-10-2013-0042.

[3] P. E. D. Love, P. Teo, and J. Morrison, “Revisiting Quality Failure Costs in Construction,” *J. Constr. Eng. Manag.*, vol. 144, no. 2, Feb. 2018, doi: 10.1061/(asce)co.1943-7862.0001427.

[4] S. D. Datta, “Artificial Intelligence and Machine Learning Applications in the Project Lifecycle of the Construction Industry: A Comprehensive Review,” *Artificial Intelligence and Machine Learning Applications in the Project Lifecycle of the Construction Industry: A Comprehensive Review*, vol. 10, 2024.

[5] S. Al-Gahtani, “ANN Prediction Model of Final Construction Cost at an Early Stage,” *Journal of Asian Architecture and Building Engineering*, vol. 24, no. 2, pp. 775–799, 2025.

[6] M. O. Sanni-Anibire, “Machine Learning-Based Framework for Predicting Cost and Schedule Performance of Construction Projects,” *Journal of Information Technology in Construction*, vol. 26, pp. 337–359, 2021.

[7] S. Banihashemi, "Machine Learning-Integrated 5D BIM Informatics: Building Materials and Construction Cost Estimation," *Innovative Infrastructure Solutions*, vol. 7, 2022.

[8] G. H. Coffie, "Using Extreme Gradient Boosting Machine Learning and SHAP to Predict Construction Project Cost Overruns," *International Journal of Construction Management*, 2024.

[9] M. Shafaay, "Modeling Construction Cost-Overrun Risks at the Front-End Engineering Design Stage," *Journal of Asian Architecture and Building Engineering*, 2025.

[10] Y. Jezzini, "Risk Allocation Model for Price Escalations in Construction Projects Using Game Theory," *Risk Allocation Model for Price Escalations in Construction Projects Using Game Theory*, 2025.

[11] M. Aslam, Z. Gao, and G. Smith, "Exploring factors for implementing lean construction for rapid initial successes in construction," *J. Clean. Prod.*, vol. 277, Dec. 2020, doi: 10.1016/j.jclepro.2020.123295.

[12] Z. Qadeer, N. A. Memon, M. A. Leghari, A. H. Memon, and F. H. Siddiqui, "Identification of Critical Factors Influencing Cost Escalation In Public Sector Construction Projects," 2019.

[13] C. H. M. P. and C. A. H. H. Yoon, "Construction Cost Index Forecasting through Multivariate Time-Series Modeling and Leading Indicators," *Journal of the Architectural Institute of Korea*, vol. 41, no. 5, pp. 319–329, 2025.

[14] G. Coffie and S. Cudjoe, "International Journal of Construction Management ISSN: (Print) (Using extreme gradient boosting (XGBoost) machine learning to predict construction cost overruns Using extreme gradient boosting (XGBoost) machine learning to predict construction cost overruns," Jul. 2023. doi: 10.1080/15623599.2023.2289754.

[15] "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *BMJ, 372, n71*, pp. 1–72, 2021.

[16] M. G. Meharie, Z. C. A. Gariy, R. N. Ndisya Mutuku, and W. J. Mengesha, "An Effective Approach to Input Variable Selection for Preliminary Cost Estimation of Construction Projects," *Advances in Civil Engineering*, vol. 2019, 2019, doi: 10.1155/2019/4092549.

[17] C.-H. Huang and S.-H. Hsieh, "Predicting BIM Labor Cost with Random Forest and Simple Linear Regression," *Autom. Constr.*, vol. 18, 2020.

[18] D. Chakraborty, H. Elhegazy, H. Elzarka, and L. Gutierrez, "A novel construction cost prediction model using hybrid natural and light gradient boosting," *Advanced Engineering Informatics*, vol. 46, p. 101201, 2020, doi: https://doi.org/10.1016/j.aei.2020.101201.

[19] P. Karadimos and L. Anthopoulos, "A Taxonomy of Machine Learning Techniques for Construction Cost Estimation," *Innovative Infrastructure Solutions*, vol. 9, 2024.

[20] F. Xiao, X. Chen, S. Yang, and J. Cheng, "Bi-objective pavement maintenance and rehabilitation optimization decision-making model incorporating the construction length of preventive maintenance projects," *Structure and Infrastructure Engineering*, vol. 21, no. 1, pp. 24–38, 2025.

[21] H. H. Elmousalami, "Artificial Intelligence and Parametric Construction Cost Estimate Modeling: State-of-the-Art Review," *J. Constr. Eng. Manag.*, vol. 146, no. 1, Jan. 2020, doi: 10.1061/(asce)co.1943-7862.0001678.

[22] J. Cranefield and P. Yoong, "The Electronic Journal for Virtual Organizations and Networks Organisational Factors Affecting Inter-Organisational Knowledge Transfer In The New Zealand State Sector: A Case Study," 2005.