,

A MBSE probabilistic framework for preliminary lifecycle costing of mechanical products

Jean-Loup Loyer

Instituto Superior Técnico, Universidade de Lisboa, Portugal

jean-loup.loyer@tecnico.ulisboa.pt

INCOSE International Symposium - Las Vegas
June 30, 2014

Table of Contents

About

Industrial Management Group within IST's Department of Mechanical Engineering

- Team of circa 10 people working on lean manufacturing, cost methods, process modeling, etc.
- Several sectors of expertise: mould-making, automotive, aeronautics, etc.
- Regular interactions with Computer Science and Mathematics Departments

Presenter's short resume

- Aerospace engineer (MSc @ Supaéro Toulouse & Imperial College London, 2007)
- Statistician and data enthusiast (MSc Statistics @ University of Toulouse, 2012)
- Diversified professional positions

About

Industrial Management Group within IST's Department of Mechanical Engineering

- Team of circa 10 people working on lean manufacturing, cost methods, process modeling, etc.
- Several sectors of expertise: mould-making, automotive, aeronautics, etc.
- Regular interactions with Computer Science and Mathematics Departments

Presenter's short resume

- Aerospace engineer (MSc @ Supaéro Toulouse & Imperial College London, 2007)
- Statistician and data enthusiast (MSc Statistics @ University of Toulouse, 2012)
- Diversified professional positions

Context and motivation

Industrial context

- Competitive pressure, challenging customer requirements, demanding airworthiness regulations
- Development of quality and profitable products
- Trading lifecycle cost (LCC) with performance, weight, quality, etc.

Importance of LCC modeling in early design of jet engines [Haskins, 2010]

- Major part of the cost typically committed during this early phase [Barton, 2001]
- Multi-attributes decision problem requiring a Systems Engineering (SE) approach
- Ability to leverage large volumes of historical data to predict future LCC

Context and motivation

Industrial context

- Competitive pressure, challenging customer requirements, demanding airworthiness regulations
- Development of quality and profitable products
- Trading lifecycle cost (LCC) with performance, weight, quality, etc.

Importance of LCC modeling in early design of jet engines [Haskins, 2010]

- Major part of the cost typically committed during this early phase [Barton, 2001]
- Multi-attributes decision problem requiring a Systems Engineering (SE) approach
- Ability to leverage large volumes of historical data to predict future LCC

Industrial case study

Rolls-Royce's jet engines

- Working principle: suck-squash-bang-woosh
- Complex mechanical product: thousands of parts, exotic alloys, unique manufacturing processes
- Modular architecture facilitating LCC estimation

MBSE framework applied to compressor blades

- Engine components adapted to probabilistic approach
  - Many items in the engine (statistical significance)
  - Subject to diverse operating conditions
  - Made of different alloys (titanium, steel, nickel)
  - Variations around generic commonality
- MBSE framework extendable to virtually any mechanical product

Industrial case study

Rolls-Royce's jet engines

- Working principle: suck-squash-bang-woosh
- Complex mechanical product: thousands of parts, exotic alloys, unique manufacturing processes
- Modular architecture facilitating LCC estimation

MBSE framework applied to compressor blades

- Engine components adapted to probabilistic approach
  - Many items in the engine (statistical significance)
  - Subject to diverse operating conditions
  - Made of different alloys (titanium, steel, nickel)
  - Variations around generic commonality
- MBSE framework extendable to virtually any mechanical product

Definition of LCC

NASA's definition of LCC

« (LCC is) the total of the direct, indirect, recurring, nonrecurring, and other related expenses incurred, or estimated to be incurred, in the design, development, verification, production, operation, maintenance, support, and disposal of a project. » [NASA, 2010]

Categories of (LCC) modeling approaches [Asiedu, 1998]

- Parametric (regression on a few variables)
- Analogous (extrapolation from similar items)
- Analytical (detailed, based on activity or physics)

Challenges in LCC estimation in early design [Meisl, 1988]

- Insufficient details in requirements or contextual parameters
- Lack of quality data

Definition of LCC

NASA's definition of LCC

« (LCC is) the total of the direct, indirect, recurring, nonrecurring, and other related expenses incurred, or estimated to be incurred, in the design, development, verification, production, operation, maintenance, support, and disposal of a project. » [NASA, 2010]

Categories of (LCC) modeling approaches [Asiedu, 1998]

- Parametric (regression on a few variables)
- Analogous (extrapolation from similar items)
- Analytical (detailed, based on activity or physics)

Challenges in LCC estimation in early design [Meisl, 1988]

- Insufficient details in requirements or contextual parameters
- Lack of quality data

Definition of LCC

NASA's definition of LCC

« (LCC is) the total of the direct, indirect, recurring, nonrecurring, and other related expenses incurred, or estimated to be incurred, in the design, development, verification, production, operation, maintenance, support, and disposal of a project. » [NASA, 2010]

Categories of (LCC) modeling approaches [Asiedu, 1998]

- Parametric (regression on a few variables)
- Analogous (extrapolation from similar items)
- Analytical (detailed, based on activity or physics)

Challenges in LCC estimation in early design [Meisl, 1988]

- Insufficient details in requirements or contextual parameters
- Lack of quality data

Current LCC approaches

Generic MBSE frameworks for LCC

- Process-based parametric (SEER-MFG, PRICE-H/HL)
- Parametric for SE-related activities (COSYSMO)
- Cost and efforts in software engineering (COCOMO II)
- Regression models for mechanical products (MISERLY)

MBSE frameworks of LCC targeted at early design

- Hybrid Activity-Based and Machine Learning applied to pressure vessels [Liu, 2008]
- Artificial Neural Networks applied to 150 products [Seo, 2002]

Limits of existing MBSE frameworks for LCC

- Often not based on actual data
- No probabilistic aspect

Current LCC approaches

Generic MBSE frameworks for LCC

- Process-based parametric (SEER-MFG, PRICE-H/HL)
- Parametric for SE-related activities (COSYSMO)
- Cost and efforts in software engineering (COCOMO II)
- Regression models for mechanical products (MISERLY)

MBSE frameworks of LCC targeted at early design

- Hybrid Activity-Based and Machine Learning applied to pressure vessels [Liu, 2008]
- Artificial Neural Networks applied to 150 products [Seo, 2002]

Limits of existing MBSE frameworks for LCC

- Often not based on actual data
- No probabilistic aspect

Current LCC approaches

Generic MBSE frameworks for LCC

- Process-based parametric (SEER-MFG, PRICE-H/HL)
- Parametric for SE-related activities (COSYSMO)
- Cost and efforts in software engineering (COCOMO II)
- Regression models for mechanical products (MISERLY)

MBSE frameworks of LCC targeted at early design

- Hybrid Activity-Based and Machine Learning applied to pressure vessels [Liu, 2008]
- Artificial Neural Networks applied to 150 products [Seo, 2002]

Limits of existing MBSE frameworks for LCC

- Often not based on actual data
- No probabilistic aspect

Table of Contents

Expression of LCC

Two-terms (simplified) equation of LCC

For a given mechanical component, at time t_0:
$LCC_{t_0} = MC_{t_0} + \sum_{i^{th} shop\ visit} R_{t_i} x MC_{t_i}$ where:

- MC_t: Manufacturing cost at a given time t
- R_{t_i}: Replacement rate at the i^{th} maintenance visit

Main features of the LCC model

- LCC as a function of time:
  - Manufacturing: materials rates, quantity, productivity...
  - Maintenance: engine deterioration, fleet structure...
- Simplified equation of LCC
  - No R&D (fuzzy) or disposal costs (negligible)
  - Manufacturing cost: no labor or consumables considered
  - Maintenance cost: repair not considered

Expression of LCC

Two-terms (simplified) equation of LCC

For a given mechanical component, at time t_0:

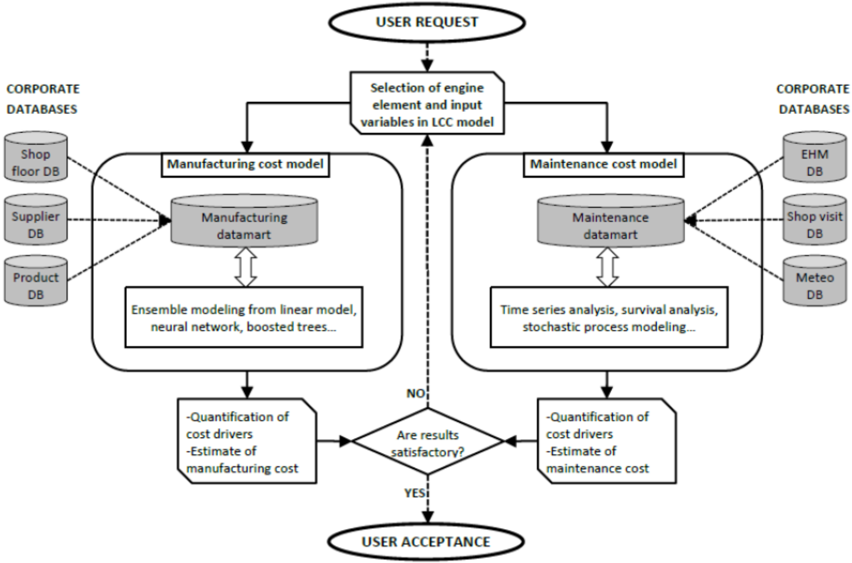
$LCC_{t_0} = MC_{t_0} + \sum_{i^{th} shop\ visit} R_{t_i} x MC_{t_i}$ where:

- MC_t: Manufacturing cost at a given time t
- R_{t_i}: Replacement rate at the i^{th} maintenance visit

Main features of the LCC model

- LCC as a function of time:
  - Manufacturing: materials rates, quantity, productivity...
  - Maintenance: engine deterioration, fleet structure...
- Simplified equation of LCC
  - No R&D (fuzzy) or disposal costs (negligible)
  - Manufacturing cost: no labor or consumables considered
  - Maintenance cost: repair not considered

Architecture of MBSE LCC framework



Royce

Estimation of manufacturing cost

Probabilistic equation of manufacturing cost

For a given mechanical component, at time t:
$MC_t = f(L, W, P, A, M, Q_t, CR_t)$ where:

- MC_t: Manufacturing cost (continuous output variable)
- $f(.)$: Complex nonlinear function to estimate (non time dependent)

Important preliminary remarks

- Model at the component level
- Time dimension due to economic-related variables
- Probabilistic aspect due to variability in the input variables

Estimation of manufacturing cost

Probabilistic equation of manufacturing cost

For a given mechanical component, at time t:
$MC_t = f(L, W, P, A, M, Q_t, CR_t)$ where:

- MC_t: Manufacturing cost (continuous output variable)
- $f(.)$: Complex nonlinear function to estimate (non time dependent)

Important preliminary remarks

- Model at the component level
- Time dimension due to economic-related variables
- Probabilistic aspect due to variability in the input variables

Characteristics of the manufacturing cost model

Seven input variables depending on each mechanical component

- Technical input variables
  - Component length L and width W
  - Manufacturing process P
  - Alloy type A and machinability M
- Economics-related input variables
  - Accumulated production volume Q_t
  - Materials cost rate CR_t

Aggregation of complementary models (Ensemble Models)

Model	Advantages	Drawbacks
Multiple Linear Regression (MLR)	Simple, interpretable	Inaccurate (linear only)
Neural Networks	Accurate (nonlinear)	Training and interpretation («black box» models)
Boosted trees	Very accurate (highly non parametric) Easy to train	Less interpretable than MLR

Characteristics of the manufacturing cost model

Seven input variables depending on each mechanical component

- Technical input variables
  - Component length L and width W
  - Manufacturing process P
  - Alloy type A and machinability M
- Economics-related input variables
  - Accumulated production volume Q_t
  - Materials cost rate CR_t

Aggregation of complementary models (Ensemble Models)

Model	**Advantages**	**Drawbacks**
Multiple Linear Regression (MLR)	Simple, interpretable	Inaccurate (linear only)
Neural Networks	Accurate (nonlinear)	Training and interpretation («black box» models)
Boosted trees	Very accurate (highly non parametric) Easy to train	Less interpretable than MLR

Results

Procedure for model assessment

- Training the model with the data on 80% of the dataset
- Evaluate the generalization (i.e. predictive accuracy) on remaining 20% of the dataset

Criteria for model assessment

- Goodness-of-fit measured by coefficient of determination R^2 (total variance of the dataset explained by the model)
- Prediction accuracy measured by percentage of good prediction on new data (i.e. held out of the training set)

Goodness-of-fit (R^2)	**Accuracy ($MAPE$)**
>0.80	>90%

Results

Procedure for model assessment

- Training the model with the data on 80% of the dataset
- Evaluate the generalization (i.e. predictive accuracy) on remaining 20% of the dataset

Criteria for model assessment

- Goodness-of-fit measured by coefficient of determination R^2 (total variance of the dataset explained by the model)
- Prediction accuracy measured by percentage of good prediction on new data (i.e. held out of the training set)

Goodness-of-fit (R^2)	**Accuracy ($MAPE$)**
>0.80	>90%

Estimation of the maintenance cost

Probabilistic equation of maintenance cost

For a given mechanical component, at time t_i:
$R_{t_i} = g(T_T^j, P_T^j, S_T^j, V_T^j, ..., AT_T, AP_T, H_T, C_T, ...)$ where:

- R_{t_i}: Average replacement rate of the part (binary output)
- $g(.)$: Complex nonlinear function to estimate
- $T = [t_{i-1}, t_i]$: Time between two maintenance visits

Important remarks

- Hundreds of input variables extracted from complex time series (air temperature AT_T or pressure AP_T...)
- Six binary classifiers used as statistical models: logistic regression, support vector machines, decision trees, random forests, boosted trees, neural networks

Estimation of the maintenance cost

Probabilistic equation of maintenance cost

For a given mechanical component, at time t_i:

$R_{t_i} = g(T_T^j, P_T^j, S_T^j, V_T^j, ..., AT_T, AP_T, H_T, C_T, ...)$ where:

- R_{t_i}: Average replacement rate of the part (binary output)
- $g(.)$: Complex nonlinear function to estimate
- $T = [t_{i-1}, t_i]$: Time between two maintenance visits

Important remarks

- Hundreds of input variables extracted from complex time series (air temperature AT_T or pressure AP_T...)
- Six binary classifiers used as statistical models: logistic regression, support vector machines, decision trees, random forests, boosted trees, neural networks

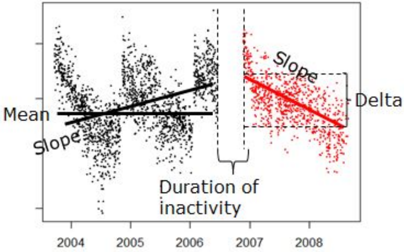
Characteristics of the maintenance cost model

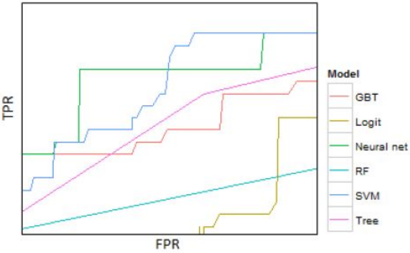
Pre-processing to extract features from raw historical sensor data

- Min/max, mean, variance
- Curve trend
- Discontinuities
- Class based on patterns
- 'Signature analysis' (wavelets, splines)

Criteria assessing model performance

- % of correctly predicted outcomes
- Area Under the ROC curve (AUC) more robust





Results

Procedure and criteria for model assessment similar to the ones used for manufacturing cost

- Training with 80% of the data and prediction accuracy on remaining 20%

Binary classifiers *in lieu* of regression models in manufacturing cost

- Goodness-of-fit measured by TPR
- Prediction accuracy measured by AUC

Goodness-of-fit (TPR)	**Accuracy (AUC)**
>85%	>85%

Results

Procedure and criteria for model assessment similar to the ones used for manufacturing cost

- Training with 80% of the data and prediction accuracy on remaining 20%

Binary classifiers *in lieu* of regression models in manufacturing cost

- Goodness-of-fit measured by *TPR*
- Prediction accuracy measured by *AUC*

Goodness-of-fit (*TPR*)	**Accuracy (*AUC*)**
>85%	>85%

Table of Contents

Take-home messages

A valid innovative framework

- Validity of the data-driven, statistics-based approach
- Scalable, reusable, maintainable model of LCC
- Contribute to structure the data within the IT system

Satisfactory results

- Complementary insights from ensemble models
- Good fit ($R^2 > 0.8 \& TPR > 85\%$) and prediction accuracy ($MAPE > 0.8 \& AUC > 85\%$) on new data, for manufacturing and maintenance costs
- Ranking of most relevant drivers LCC, challenging traditional engineering approach (results not presented today)

Take-home messages

A valid innovative framework

- Validity of the data-driven, statistics-based approach
- Scalable, reusable, maintainable model of LCC
- Contribute to structure the data within the IT system

Satisfactory results

- Complementary insights from ensemble models
- Good fit ($R^2 > 0.8 \& TPR > 85\%$) and prediction accuracy ($MAPE > 0.8 \& AUC > 85\%$) on new data, for manufacturing and maintenance costs
- Ranking of most relevant drivers LCC, challenging traditional engineering approach (results not presented today)

Benefits and limits of the MBSE framework

Benefits

- Generic framework
  - Applicable at system, subsystem and component levels
  - Applicable to many (mechanical) products and sectors
- Probabilistic, data-driven approach
  - Risk in decision-making and financial evaluation
  - Estimation based on actual data (no dubious assumptions needed)

Limits

- Data-driven (!)
  - Requires sufficient amount of actual data
  - Pre-existing databases and data-rich IT infrastructure
  - Lack of statistical training amongst engineers
- Less applicable to families of products with low commonality

Benefits and limits of the MBSE framework

Benefits

- Generic framework
  - Applicable at system, subsystem and component levels
  - Applicable to many (mechanical) products and sectors
- Probabilistic, data-driven approach
  - Risk in decision-making and financial evaluation
  - Estimation based on actual data (no dubious assumptions needed)

Limits

- Data-driven (!)
  - Requires sufficient amount of actual data
  - Pre-existing databases and data-rich IT infrastructure
  - Lack of statistical training amongst engineers
- Less applicable to families of products with low commonality

What remains to be done

Refine the models

- Add more input variables, better extracted from time series (maintenance cost)
- Use more sophisticated models (Dynamic Bayesian networks, Functional Data Analysis, Markov-like models...)
- Interactions between drivers of manufacturing and maintenance costs

Benchmark and extend the results

- Compare with accuracy and variable importance obtained from current «manual» estimates of LCC by Rolls-Royce cost engineers
- Verify validity on different components or sectors

What remains to be done

Refine the models

- Add more input variables, better extracted from time series (maintenance cost)
- Use more sophisticated models (Dynamic Bayesian networks, Functional Data Analysis, Markov-like models...)
- Interactions between drivers of manufacturing and maintenance costs

Benchmark and extend the results

- Compare with accuracy and variable importance obtained from current «manual» estimates of LCC by Rolls-Royce cost engineers
- Verify validity on different components or sectors

References

Asiedu Y. and P. Gu. 1998.
Product life cycle cost analysis: state of the art review. *International Journal of Production Research*. 36(4): 883–908

Barton J. A. , D. M. Love and G. D. Taylor. 2001.
Design determines 70% of cost? A review of implications for design evaluation. *Journal of Engineering Design* 12 (1): 47-58.

Haskins, C. 2010.
Systems Engineering Handbook: A Guide for System Life Cycle Processes and Activities. Version 3.2. Revised by M. Krueger, D. Walden, and R. D. Hamelin. *San Diego, CA (US): INCOSE*

Liu H., V. Gopalkrishnan, W.K. Ng, B. Song and X. Li. 2008.
An intelligent system for estimating full product Life Cycle Cost at the early design stage. *International Journal of Product Lifecycle Management* 2 (2-3): 96-113

Meisl C.J. 1988.
Techniques for cost estimating in early program phases. *Engineering Costs and Production Economics* 14 (2), 95 – 106.

NASA. 2008.
NASA Cost Estimating Handbook. *Washington, DC (US): National Aeronautics and Space Administration* NASA Headquarters - Cost Analysis Division.

Seo K.-K., J.-H. Park, D.-S. Jang and D. Wallace. 2002.
Approximate Estimation of the Product Life Cycle Cost Using Artificial Neural Networks in Conceptual Design. *The International Journal of Advanced Manufacturing Technology*. 19 (6): 461-471