Introduction to LS-OPT

The Arup Campus, Blythe Gate, Blythe Valley Park, Solihull, West Midlands, B90 8AE
tel: +44 (0) 121 213 3399
e-mail: dyna.support@arup.com

12th September 2019
1. Introduction
2. Design Formulation
3. Optimisation Methods
4. Point Selection
5. Optimisation Strategy
6. Viewer
7. Sensitivity Analysis
8. Error Analysis
9. Multi-Objective Optimisation
10. Parameter Identification
11. Conclusion
Examples

http://www.lsoptsupport.com/documents
Introduction
**What is Optimisation?**

**Optimisation**: A procedure for achieving the best outcome of a given operation while satisfying certain restrictions.

Infeasible design space due to design constraints.
**LS-OPT** : A stand-alone design optimisation & probabilistic analysis tool with an interface to LS-DYNA (and other solvers).
LS-OPT GUI

LS-OPT Viewer
- optimisation results

LS-OPT Flowchart
- optimisation process
LS-OPT Applications

Multi-objective optimisation
Optimising a design with respect to multiple objectives.
*e.g. Minimising passenger injury response with respect to crush can thickness and length.*

Worst-case design
Optimising a design for the worst case (minimising a response with respect to some variables while maximising the same response with respect to other variables).
*e.g. Minimising passenger injury response with respect to crush can design while maximising passenger injury response with respect to angle of impact.*

Reliability based design
Optimising a design with consideration of uncertainties in variables and responses.
*e.g. Minimising passenger injury response with respect to crush can design including consideration of variations in material properties and manufacturing tolerances.*

Parameter identification
Calibration of a calculated curve against an input target curve.
Any optimisation problem consists of three key components:

**Design Variables**
Design parameters which may be changed to produce different design responses.

\[ \mathbf{x} = [x_1, x_2, x_3, \ldots, x_n] \]

For \( n \) variables

- e.g. Section area
- Component length
- Material grade

**Design Objectives**
An optimum target for one or more of these design responses.

\[ \min [f_i(\mathbf{x})] \]

For \( N \) objectives

- e.g. Minimise injury
- Minimise mass/cost
- Maximise energy absorption

**Design Constraints**
Restrictions on the design parameters and/or design responses.

\[ L_j \leq g_j(\mathbf{x}) \leq U_j \]

For \( M \) constraints

- e.g. Stress < 50 MPa
- Displacement < 50mm
- 1.7 Hz < Frequency < 1.9 Hz
- Cost & Mass
Parameter Setup

LS-OPT accepts a number of design variable types such as constants, strings, dependent variables and noise variables (which include a probability distribution). This course focuses on two types:

- Continuous variables (e.g. values between 1 and 5)
- Discrete variables (e.g. values 1, 2, 3, 4 and 5)

It is recommended that continuous variables are used, except where an integer/binary input is essential e.g. LS-DYNA material property reference number, on/off input for a support condition.
Solver Setup

The design variables are then processed by a solver (e.g. LS-DYNA, Excel, Perl code) to produce responses.
Histories and Responses

In addition, the user specifies which outputs to monitor:

1. **Responses** (solver output at a particular timestep)
2. **Histories** (solver output over numerous timesteps)
Composites
Composites are used to combine response surfaces into a single expression. They are useful in:
1. Specifying objectives
2. Specifying constraints
3. Viewing results

Composites are also used with responses and histories for the purpose of parameter identification, to be discussed later.

Expression composite, used to combine number of responses e.g. distance between two nodes

Standard Composite and Curve Matching, used for parameter identification
Objectives and constraints

Objectives and constraints are specified by selecting from a list of responses and composites.

1. **Objective function**: Minimisation (default) or maximisation of the selected response/composite.
2. **Objective weighting**: Equal weighting (default) or varying with relative importance of each objective.
3. **Constraint upper/lower bound**
4. **Constraint internal scaling**: Normalisation of constraints to eliminate large variations in the magnitudes of design variables and constraints.

*Constraint internal scaling should always be used if you have more than one constraint (if scaling is turned off, constraint violations are absolute values which can mean very different things from one constraint to the next).*
Demonstration

Simple optimisation

N432 (Disp2)

N167 (Disp1)

Intrusion (Disp1 – Disp2)
Optimisation Methods
1. **Variables**
   - Setup
     - 2 parameters
   - Finish

2. **Responses**
   - Sampling 1
     - 2 vars, 5 d-opt designs
   - 1
     - 2 pars, 5 resps

3. **Composites**
   - Composites
     - 1 definition
   - Build Metamodels
     - 5 linear surfaces

4. **Objectives & constraints**
   - Optimization
     - 1 objective
     - 2 constraints

5. **Optimisation method**
   - Domain reduction (SRSM)
     - 10 iterations
Three key optimisation methods are utilised in LS-OPT.

- **Gradient Based Methods**
  - Sensitivity (LS-OPT categorises this as a metamodel based method)

- **Metamodel (Response Surface) Based Methods**
  - Response Surface Method (RSM)
  - Successive Response Surface Method (SRSM)
  - Polynomial Methods
  - Neural Network Methods

- **Direct Optimisation Methods**
  - Genetic Algorithm
A response surface is an approximation of the design response. Once the response surface is built, an optimisation algorithm such as a gradient based algorithm or a genetic algorithm.

**What is the difference?**

**Response**
- A numerical indicator of the performance of the design.

**Response surface**
- A mathematical expression which relates the design variables to design response.
- Also known as metamodels, surrogate models, proxy models, and representative models.

**Advantages**
- Can be applied to highly non-linear responses
- Can be applied for noisy response
- Does not require analytical design sensitivities
- Displays a tendency to capture a global optima
Metamodel Based Methods – Polynomial

Polynomial order can be linear, linear with interaction, quadratic or elliptic.
Metamodel Based Methods – Neural Networks

Neural networks consist of an input layer, one or more hidden layers (nodes), and an output layer, each containing a number of neurons.

\[ x = (x_1, x_2, \ldots, x_K) \]

\[ y(x, W) \]

Weights & biases of hidden layer

Weights & bias of output layer
Summary – Metamodels

**Polynomial**
- Very robust, especially for sequential optimisation methods
- The user is obliged to choose the order of the polynomial
- Greater possibility for a bias error of a nonlinear response

**Neural Network**
- Good approximations, even when not so many design points are used.
- Can be applied to highly nonlinear responses
- Can be applied to noisy responses
- Global accuracy, can be used for design space exploration and studying trade-off (multi-objective optimisation)

**First-order**
- Inexpensive
- Useful within a certain sub-region
- May require 5 to 10 iterations for convergence
- Poor global accuracy

**Second order**
- More robust, due to curvature consideration
- Global accuracy
- Can be more expensive

There is a good overview on metamodels available here:

https://www.lsoptsupport.com/faqs/setup/which-meta-model
Point Selection
Point selection (experimental design or DOE) is an important component of the metamodel based optimisation methods.

**Design space:** A region in the n-dimensional space of the design variables \((x_1, x_2, \ldots x_n)\) to which the design is limited.

**Point selection:** The method by which particular design variable combinations (points within the design space) are selected for testing.
The point selection methods available in LS-OPT are dependent on the metamodel.

![Point Selection Method Options](image)
Full Factorial

Full Factorial (First Order) $\rightarrow 2^n$ points
($n$ – number of design variables)

$n = 2$

$x_1$
$x_2$

$n = 3$

$x_1$
$x_2$
$x_3$

Full Factorial (Second Order) $\rightarrow 3^n$ points
($n$ – number of design variables)

$n = 2$

$x_1$
$x_2$

$n = 3$

$x_1$
$x_2$
$x_3$
Linear Koshal (First Order) → \( n + 1 \) points
\((n – \text{number of design variables})\)

Quadratic Koshal (Second Order) → \( 1 + n + n(n + 1)/2 \) points
\((n – \text{number of design variables})\)
D–Optimal

**D-Optimal (Determinant optimal)**

D-Optimal selects the best distribution of a specified number of design points $p$, by maximising the determinant $|X^T X|$ where $X$ is a $n \times p$ matrix containing all design points.

The design points are selected from a **basis set**, which is determined using another point selection method. The default point selection method is based on the number of variables $n$.

- Small values of $n$ → Full Factorial
- Large values of $n$ → Space filling

For $n > 50$, Latin Hypercube is much faster.

**Recommended point selection for polynomial metamodels**

e.g. $p = 10$, $n = 2$, basis set fifth order full factorial ($5^n$ points)
Space Filling
For a specified number of design points $p$, the space filling algorithm maximises the minimum distance between points.

Latin Hypercube
Latin Hypercube Sampling (LHS) is a constrained random experimental design.

For a design space with $n$ design variables, the range of each variable is divided into $m$ equally probable intervals. For a specified number of design points $p$, each point must only contain one variable per hyperplane, (thus $m!^{(n-1)}$ is the maximum number of combinations).

e.g. $n = 2$, $m = 3$, $p = 6$

$3!^{(2-1)} = 6$ Latin Hypercube combinations
Optimisation Strategies
1. Variables
   - Setup
     - 2 parameters

2. Responses
   - Point selection
     - Sampling 1
       - 2 vars, 5 d-opt designs
   - 1
     - 2 pars, 5 resps

3. Composites
   - Build Metamodels
     - 5 linear surfaces

4. Objectives & constraints
   - Optimization
     - 1 objective
       - 1 definition
     - 2 constraints

5. Optimisation method

6. Optimisation strategy
   - Finish
   - Domain reduction (SRSM)
     - Termination criteria
       - 10 iterations

Slide 33
There are 3 optimisation strategies in LS-OPT:

- **Single Iteration**
  - Sampling and optimisation are done in a single iteration

- **Sequential**
  - Sampling and optimisation are added sequentially in the full design space

- **Sequential with Domain Reduction**
  - Sampling and optimisation are added sequentially in an adaptive subregion
  - Metamodel optimisation is done at each iteration and is limited to the current subregion
The strategy adopted depends on whether you want to:

- Find an optimal set of parameters or explore globally
- Solve the problem iteratively or adhere to a specified ‘simulation budget’
Single Iteration

Application
• Suitable for global exploration.

Metamodel
• Suitable for neural networks.

Point Selection
• Choose a large number of points as there is only one iteration.
• For neural networks, choose 'space filling' points.
**Sequential Optimisation strategy**

**Application**
- Suitable for global exploration.

**Metamodel**
- Suitable for neural networks.

**Point Selection**
- Choose a small number of points for each iteration. i.e. \([1.5(n+1)+1]\) where ‘n’ is the number of variables.
- For neural networks, choose ‘space filling’ points. Note points belonging to previous iterations are included.

**Accuracy**
- For a large number of experimental points accuracy is similar to single iteration.
- Optimisation can be stopped depending on accuracy (more flexible than single iteration).
Sequential with Domain Reduction

Application
- Ideal for system identification.
- Not suitable for global exploration.

Metamodel
- Suitable for polynomials (Linear recommended).
- Suitable for neural networks.

Point Selection
- Choose a small number of points for each iteration.
  i.e. $[1.5(n+1)+1]$ where ‘n’ is the number of variables.
- For polynomials, choose ‘D-Optimal’ points.
  Note points belonging to previous iterations are ignored.
- For neural networks, choose ‘space filling’ points.
  Note points belonging to previous iterations are included.

Accuracy
- For a large number of experimental points accuracy is similar to single iteration.
- Optimisation can be stopped depending on accuracy (more flexible than single iteration).

First Order Polynomials
Inexpensive ($\propto n$)
Robust iterative method
Note cycling (oscillation) can occur

Second Order Polynomials
Expensive (Quadratic $\propto n^2$ Elliptical $\propto 2n$)
More accurate. Better for trade-off studies
Convergence Criteria

Design change tolerance

Error norm for design variables

\[ \epsilon_x = \frac{\|x^{(k)} - x^{(k-1)}\|}{\|x^{(k)}\|} \]

Objective function tolerance

Error norm for objective function

\[ \epsilon = \frac{\|f^{(k)} - f^{(k-1)}\|}{\|f^{(k)}\|} \]

Response accuracy tolerance

Error norm of SPRESS/mean of response \( i \)

\[ \epsilon = max|s_i^{(k)} - s_i^{(k-1)}| \]

- Can use **AND** (all) or **OR** (whichever occurs first) for the above three tolerance measures.
- The optimisation **STOPS** when tolerance is satisfied or when the number of iterations specified by the user have been completed – whichever is reached first.
LS-OPT Viewer
LS-OPT Viewer offers a number of graphical plot options to view the results. We will now take a look at some key plots.

**Simulations**
Focuses on the computed data (e.g. LS-DYNA results)

**Metamodel**
Focuses on the metamodel predicted data

**Optimisation**
Focuses on the genetic algorithm or metamodel predicted data over a number of iterations
Correlation Matrix, Scatter Plots & Statistical Tools

Lower left cells provide scatter plots (e.g. HIC vs thood).

Upper right cells provide the respective $R^2$ coefficient of multiple determination (e.g. HIC vs thood = 0.83).

Diagonal cells provide statistical plots indicating the number of samples, mean and standard deviation for each variable, response or composite.
Parallel Coordinates

Maximum and minimum constraints on any of the variables/responses/composites can be varied using the sliders.

Each line represents a particular data point. Blue lines are feasible (sit between the two red lines), grey lines are infeasible.

Data values for a particular data point where HIC is minimum (while satisfying Intrusion < 500mm).
Surface plot allows us to view the response surface in 3D.

Computed data points may also be plotted, with feasible points coloured green, and infeasible points coloured red.
2D Interpolator

2D interpolator allows us to view a ‘slice’ of the response surface in 2D. This is particularly useful for multi-dimensional optimisation problem.
Sensitivity Analysis
Sensitivity Analysis

Sensitivity analysis allows the user to identify and remove variables which have insignificant contributions to the metamodel.

→ Reduce number of variables.
→ Reduce computational effort.

There are two sensitivity analysis tools in LS-OPT:
- Analysis of Variance (ANOVA)
- Global Sensitivity Analysis (GSA)
Analysis of Variance (ANOVA)

- **Linear** sensitivity measure (effect of variables investigated independently, cannot detect quadratic correlation).
- Only considers a **single response**.
- Returns the influence of each variable as a positive or negative value, normalised with respect to design space.
- Automatically performed as part of a DOE study or optimisation run.

\[ b_j = \frac{\partial f}{\partial x_j} \Delta x_j \]

**Importance of variable** \( b_j \)

**Uncertainty of variable** \( \Delta b_j \)

Based on regression analysis
Global Sensitivity Analysis (GSA)

- **Non-linear** sensitivity measure (considers the interaction between several design variables as well as curvature).
- Can consider **multiple responses**.
- Returns influence of each variable as an absolute value, Sobol’s Index $0 \leq S_i \leq 1$,
  \[ S_i = \frac{\text{Variance of response due to } x_i}{\text{Total variance of response}} \]
- Computationally expensive – needs to be explicitly specified as part of a DOE study or optimisation run.

*Multiple variables*  
*Multiple responses*

- `thood (87.6% - 87.6%)`
- `tbumper (12.4% - 100.0%)`

% Influence on Responses/Composites
The ANOVA and GSA should be considered in parallel.

**ANOVA**
- The error bars represent the contribution of the variables to noise or poorness of fit in a linear approximation of the computed responses.
- The bar lengths represent the influence of the variables on the linear approximation.
- The bar directions represent whether the variables have a positive or negative influence on the linear approximation.

**GSA**
- The bar lengths represent the influence of the variables on the response surface, as a percentage of the total variance of the response surface.
- The total variation value (provided at the top of the GSA plot) indicates this total variance.
- The noise variation value (also provided at the top of the GSA plot) indicates the noise or poorness of fit of the response surface to the computed responses.
Example 1

\[ G = 0.1x_2 + 0.5 \]
Example 2

\[ F = 3x_1^2 + 2x_2 + 3 \]
Error Analysis
Error Analysis

Error analysis allows the user to understand metamodel capability of predicting the accurate response.

Accuracy is dependent on numerous factors including:

- **Size of the region of interest**
  *The smaller the size, the more accurate the surface (up to a point where the noise dominates).*

- **Number and distribution of the experimental points**
  *More points give better predictive capability.*

- **Order and nature of the approximating function**
  *Higher order is more accurate, but overfitting can occur.*
There are numerous formulas to quantify errors

**Maximum Residual**
\[
\varepsilon_{\text{max}} = \max |y_i - \hat{y}_i|
\]

**Root Mean Square**
\[
\varepsilon_{\text{RMS}} = \sqrt{\frac{1}{P} \sum_{i=1}^{P} (y_i - \hat{y}_i)^2}
\]

**R² coefficient of multiple determination**
\[
R^2 = \frac{\sum_{i=1}^{P} (\hat{y}_i - \bar{y}_i)^2}{\sum_{i=1}^{P} (y_i - \bar{y})^2}
\]

- \(P\) – number of design points
- \(y_i\) – actual/ computed response
- \(\hat{y}_i\) – predicted response
- \(\bar{y}_i\) – mean of the actual responses
- \(h_{ii}\) – diagonal terms of ‘H’
- \(H\) – matrix that maps the observed responses to the fitted responses, i.e., \(\hat{y} = Hy\)

**Prediction Sum of Squares (Leave One Out)**
Calculated by iteratively excluding point \(i\) in the least squares calculation and predicting point \(i\) from this fit.

\[
PRESS = \sum_{i=1}^{P} \left( \frac{y_i - \hat{y}_i}{1 - h_{ii}} \right)^2
\]

**Square Root of Prediction Sum of Squares**
\[
SPRESS = \sqrt{\frac{1}{P} \sum_{i=1}^{P} \left( \frac{y_i - \hat{y}_i}{1 - h_{ii}} \right)^2}
\]

**R² for Prediction**
\[
R^2_{\text{prediction}} = 1 - \frac{\text{PRESS}}{S_{yy}}
\]

\[
S_{yy} = y^T y - \frac{1}{P} \left( \sum_{i=1}^{P} y_i \right)^2
\]
Example 1

Metamodelling Accuracy
For Response Function "HIC"
Linear: RMS Err = 76.4 (27.7 %), Sqrt PRESS = 252 (91.4 %), R-sq = 0.923

Predicted Response Value vs. Computed Response Value
Metamodeling Accuracy
For Response Function "HIC"
Linear: RMS Err = 0.962 (0.402 %), Sqrt PRESS = 3.21 (1.34 %), R-sq = 0.998
Multi-Objective Optimisation
Most real-life engineering problems deal with multiple objectives (e.g., cost, weight, safety, efficiency). Often there is not a single optimal solution, but a trade-off between conflicting requirements (e.g., weight vs. efficiency).

**Pareto optimal set**
A set of non-dominated (equally optimal) solutions.
For Multi-Objective Optimisation (MOO) problems, population based methods like Genetic Algorithms (GA) are attractive because many trade-off solutions can be found in a single simulation run.

Adding multiple objectives in LS-OPT is straightforward, however selecting the optimisation method and interpreting results requires more thought.

**Direct Simulation based GA**
- Optimise without creating metamodels
- Use simulations to evaluate designs – high computational cost.
- Accurate results

**Metamodel based GA**
- Use metamodels to evaluate functions
- Computationally inexpensive
- Accuracy depends on the quality of metamodels
Interpreting Results – Trade-off Plot

Similar to the Metamodel Plot, the Trade-off Plot allows the user to view all simulation points, with one objective/response on the z-axis and a second objective/response as a colour scale.

In addition, the Trade-off plot allows the user to view only those simulation points which are Pareto optimal solutions.
Parameter identification is based on minimising the mismatch between target values and solver-computed outputs.

### LS-OPT Parameter Identification

**Parameter Identification Methods**
There are two key parameter identification methods available in LS-OPT:

1. **MSE or Ordinate Based Curve Matching**
   (Point based or History based)

2. **Curve Mapping**
   (History based)

**Optimisation Settings**
The following optimisation settings are recommended:

- Linear polynomial metamodel
- D-optimal point selection
- Sequential domain reduction
Ordinate based curve matching

Optimisation is based on the Mean Squared Error (MSE) metric which calculates the discrepancy in ordinate values between two curves.

\[
MSE = \frac{1}{P} \sum_{p=1}^{P} W_p \left[ \frac{f_p(x) - G_p}{S_p} \right]^2
\]

- \(P\) is the number of points,
- \(S_p\) is the residual scale factor (normalisation of error), and
- \(W_p\) is the weight factor.

Application

Ordinate based curve matching is used for non-steep curves with unique abscissa values only.

- Steep portions of the response are difficult or impossible to incorporate (e.g. linear elastic range or failure).
- Hysteretic test curves cannot be matched since the abscissa values are non-monotonic/non-unique.
- Results in instability if ranges of the computed and test curve do not coincide in the y-axis (ordinate) at an interim stage of the optimisation.
- Partial matching is not robust.
1. Set up a history definition for the computed curve.
2. Create a curve matching composite.
   - Set algorithm to MSE (shown) or Curve Mapping.
   - Specify the target curve in a history file and import.
   - Specify the computed curve.
3. Define the composite as an objective function.
4. Run optimisation.
Parameter identification results are often best viewed using History Plots. By displaying only the optimum solution for each iteration, and colouring the curves by iteration, we can clearly see the curves converge on the input data points.
Interpreting Results

The calculated parameter values, and the convergence of the optimisation to these values can be viewed as an optimisation History Plot.

![Optimization History](image)

- **Reduction in domain**
- **Converged solution**
Demonstration
Parameter Identification

Elastoplastic material
- Young’s modulus?
- Yield stress?

Rigid base

Applied displacement
Conclusion
Contents

1. Introduction
2. Design Formulation
3. Optimisation Methods
4. Point Selection
5. Optimisation Strategy
6. Viewer
7. Sensitivity Analysis
8. Error Analysis
9. Multi-Objective optimisation
10. Parameter Identification
11. Conclusion

LS-OPT Introduction

Slide 70
Glossary

- ANOVA – Analysis of Variance
- ASA – Adaptive Simulated Annealing
- DOE – Design of Experiments
- FFNN – Feed Forward Neural Network
- GA – Genetic Algorithm
- GSA – Global Sensitivity Analysis
- GUI – Graphical User Interface
- LFOP – Leap-Frog Optimiser
- LHS – Latin Hypercube Sampling
- MOO – Multi-Objective Optimisation
- MSE – Mean Square Error
- NN – Neural Network
- PRESS – Prediction Sum of Squares
- RBF – Radial Basis Function
- RMS – Root Mean Square
- RSM – Response Surface Method
- SRSM – Successive Response Surface Method
For more information please contact the following:

**UK Contact:**
The Arup Campus
Blythe Valley Park
Solihull
B90 8AE
United Kingdom
T +44 (0)121 213 3399
dyna.support@arup.com

**China Contact:**
Arup China
39/F-41/F Huaihai Plaza
1045 Huaihai Road (M)
Xuhui District, Shanghai
200031
China
T +86 21 6126 2875
china.support@arup.com

**India Contact**
Arup India
Ananth Info Park, HiTec City
Madhapur Phase-II
Hyderabad
500 081, Telangana
India
T +91 (0) 40 44369797 / 8
india.support@arup.com

**USA West Contact**
Arup Americas
c/o 560 Mission Street Suite 700
San Francisco
CA 94105
United States
T +1 415 940 0959
us.support@arup.com

www.arup.com/dyna