A New Multiobjective Genetic Programming for Extraction of Design Information from Non-dominated Solutions

ISAS/JAXA, Japan Tomoaki Tatsukawa
ISAS/JAXA, Japan Taku Nonomura
ISAS/JAXA, Japan Akira Oyama
ISAS/JAXA, Japan Kozo Fujii

Background

Multi-Objective Design Exploration(MODE)

Obayashi, et.al. , "Multi-Objective Design Exploration for Aerodynamic Configurations" , *2005*

Problem Setting

⇩

Multi-objective optimization

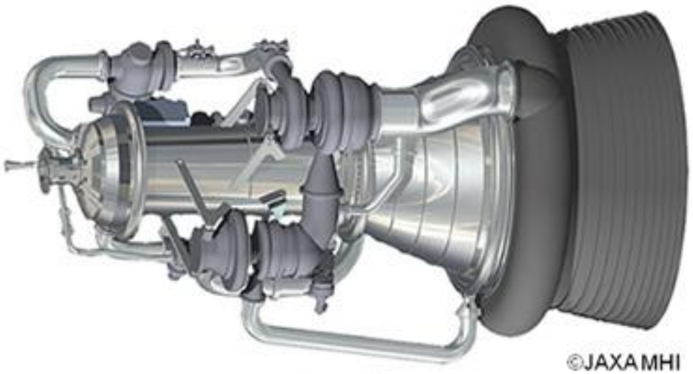
⇩

Data mining methods

⇩

Extraction of design knowledge

Mitsubishi Regional Jet (MRJ)

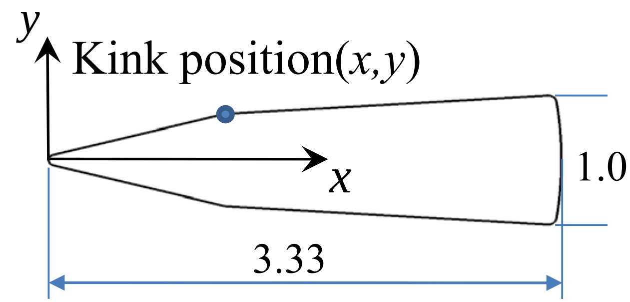


LE-X

Multi-Objective Design Exploration - Problem setting -

■ Design paramters

Kink position (x, y)



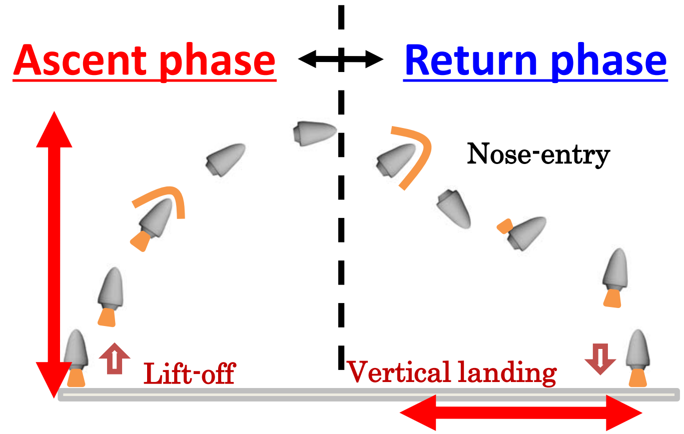
■ Objective functions

Ascent phase

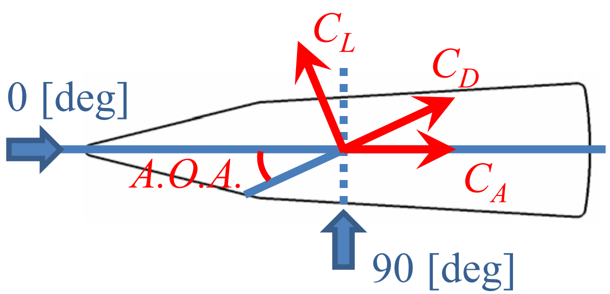
1. Drag to be minimized (M=2.0, AOA = 0)

Return phase

2. Maximum L/D to be maximized at subsonic condition (M=0.8)
3. Maximum L/D to be maximized at supersonic condition (M=2.0)



4. Body volume to be maximized



Multi-Objective Design Exploration

- Multi-objective optimization & Data mining -

Problem Setting

⇩

Multi-objective optimization

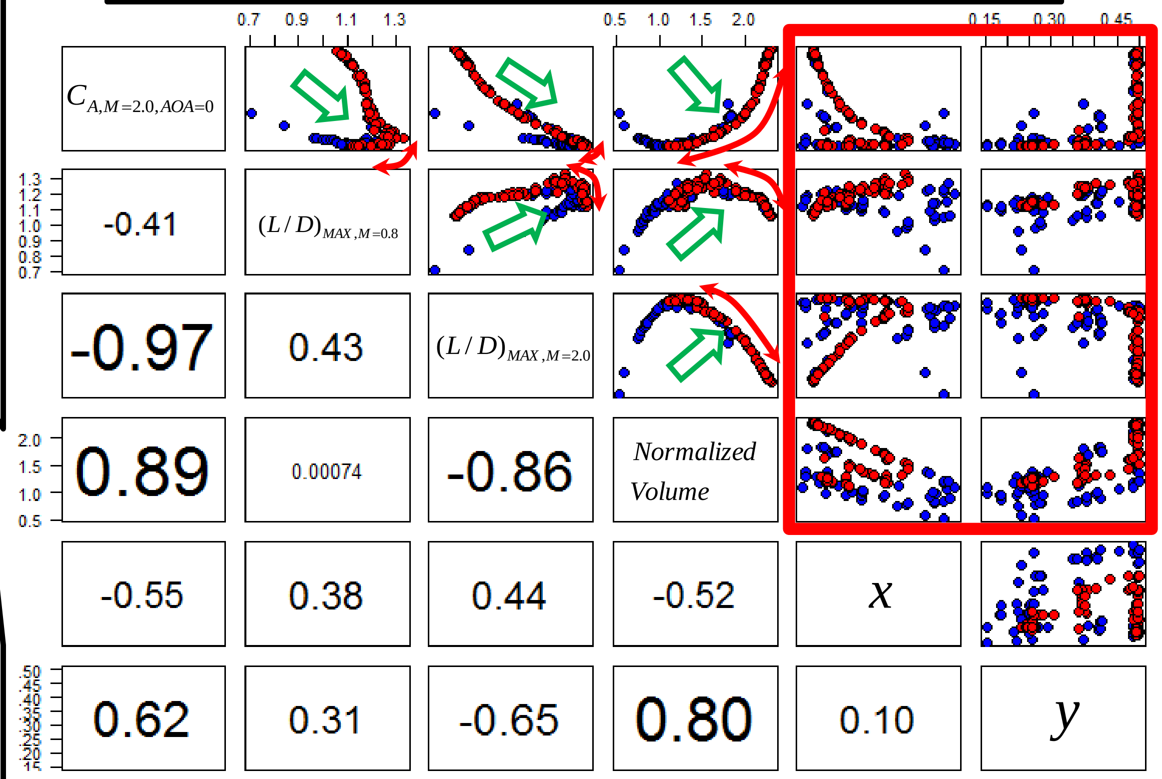
⇩

Data mining methods

⇩

Extraction of design knowledge

Scatter plot matrix with correlation coefficient



- **It is difficult to reveal non-linear relationship (sin(x1), x1x2, x1/x2) between parameters**

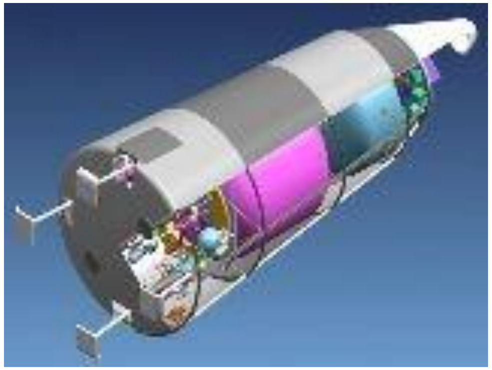
➡ Genetic programming is one of the evolutionary technique to reveal non-linear relationship.

Objective

- Present a new data mining method for MODE to extract design information from non-dominated solutions

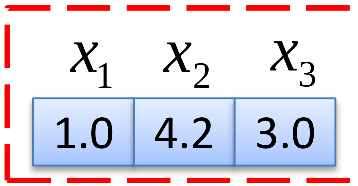
Approach

- Multi-objective GP(MOGP) for MODE is introduced
- MOGP is applied to the practical multi-objective optimization problem of reusable launch vehicle

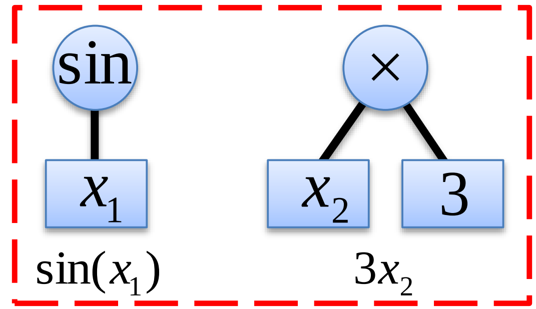


About Genetic Programming

- Extension of Genetic Algorithm(GA) (*Koza, 1990*)
- Genome expression of GP is different from GA



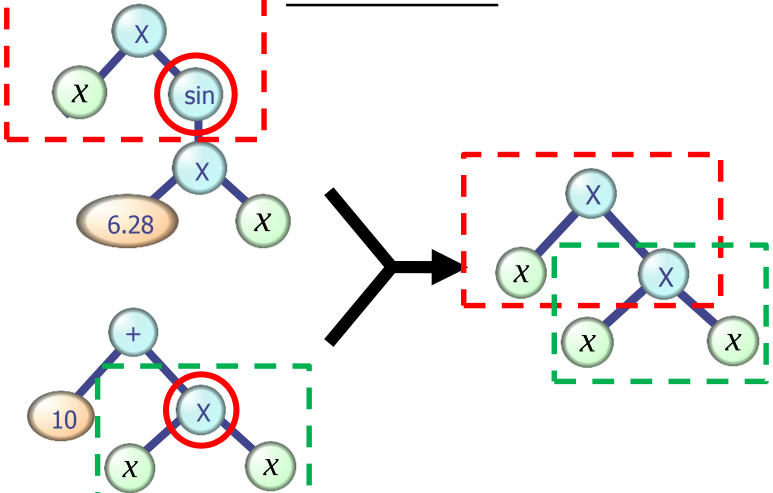
GA (Array)



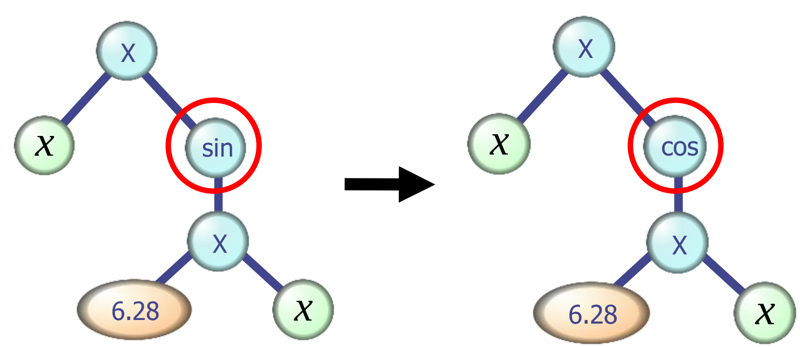
GP(Tree expression)

- Genetic operators

Crossover



Mutation



Genetic Programming in MODE

Problem Setting

Multi-objective optimization

Data mining methods

Extraction of design knowledge

GP is used as symbolic regression technique

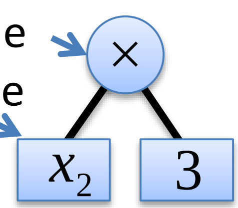
- **Extract effective non-linear terms**
- **Give us hint of unknown new parameters**

■ Objective functions

1. The measure of accuracy
Maximization of correlation coefficient

2. The measure of complexity
Minimization of the number of nodes

Function node
Terminal node



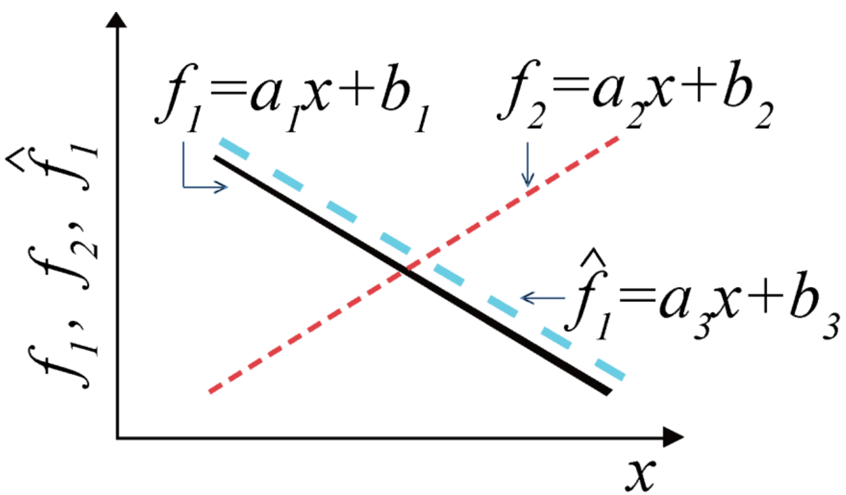
The number of node = 3

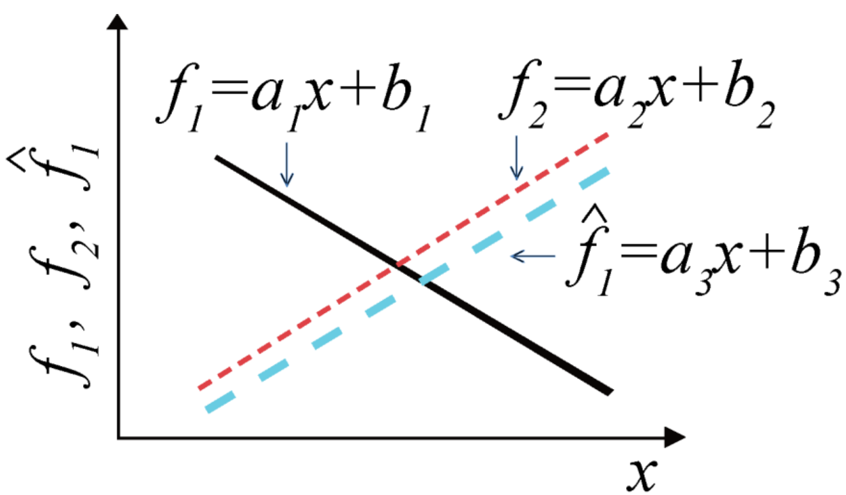
Proposed MOGP
- Measurement of Accuracy-

Proposed GP

Maximization of squared correlation coefficient

$$Cor^2(f_i, \hat{f}) = \left(\frac{\sum_{j=1}^{N} (f_{i,j} - \bar{f}_i)(\hat{f}_j - \bar{\hat{f}})}{\sqrt{(f_{i,j} - \bar{f}_i)^2} \sqrt{(\hat{f}_j - \bar{\hat{f}})^2}} \right)^2 \longrightarrow \max \qquad i = 1,2,\ldots,M$$





We want to extract “x” as design information

Proposed MOGP
- Objective Function -

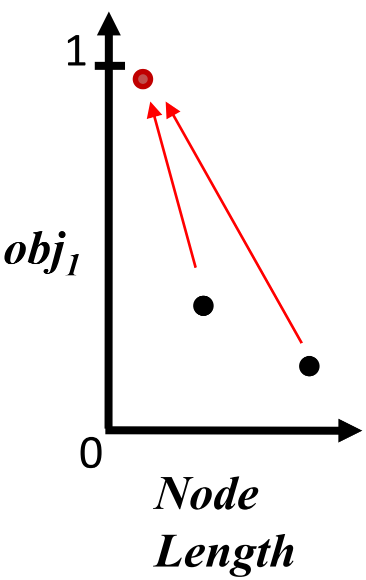
GP

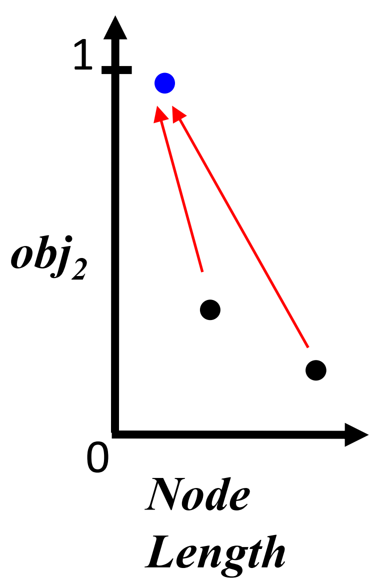
1) Maximization of correlation coefficient between obj_i
2) Minimization of the number of nodes

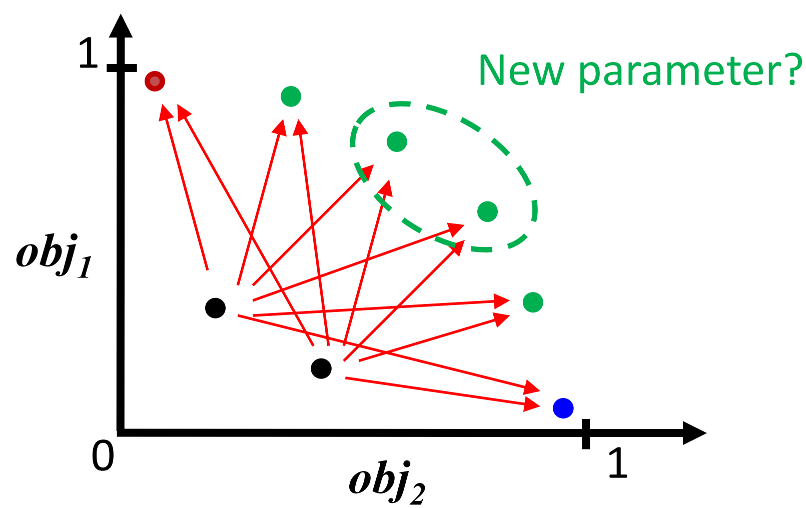
MOGP

1) Maximization of squared correlation coefficient between obj_1
2) Maximization of squared correlation coefficient between obj_2
⋮
N) Maximization of squared correlation coefficient between obj_N
N+1) Minimization of the number of nodes

obj_i : *ith* objective function of MOO



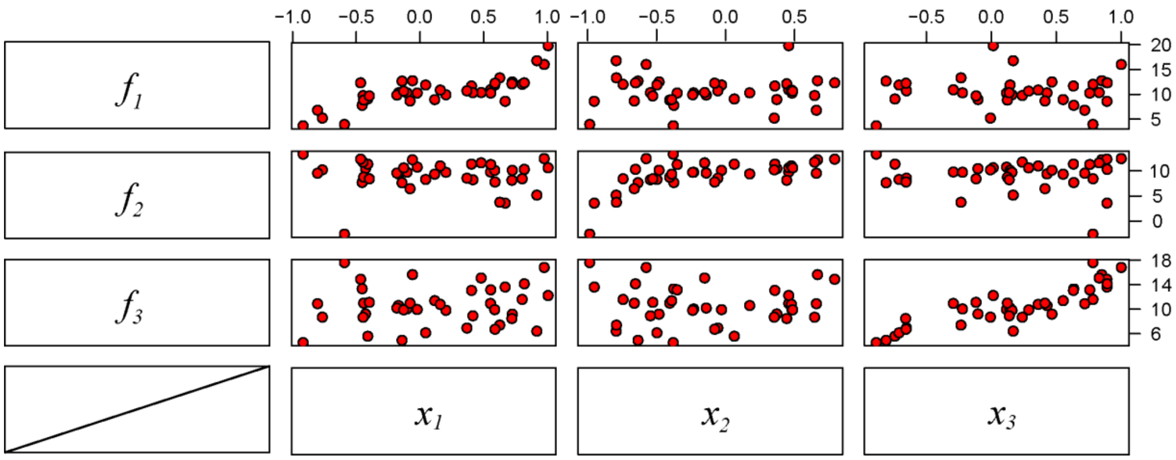




Symbolic Regression Example

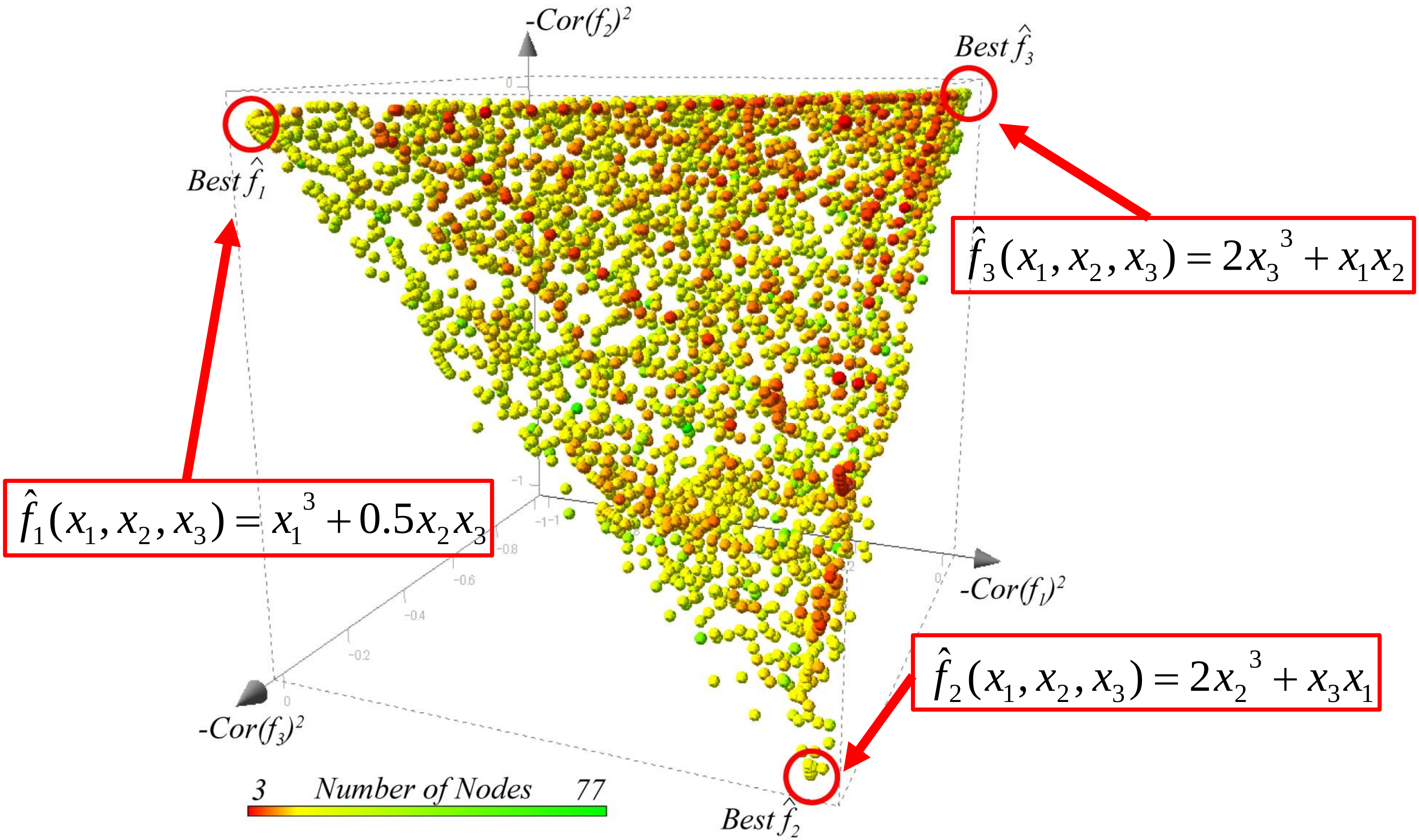
- Test function
$$f_1(x_1,x_2,x_3)=10x_1^{\ 3}+5x_2x_3+10$$
$$f_2(x_1,x_2,x_3)=10x_2^{\ 3}+5x_3x_1+10$$
$$f_3(x_1,x_2,x_3)=10x_3^{\ 3}+5x_1x_2+10 \qquad x_1,x_2,x_3=[-1,1]$$

- Data set
  - 40 sample points (random)



- GP
  - 1000 Individuals, 1000 Generations
  - 15 Trial

Results of Proposed MOGP



MODE of reusable launch vehicle (RLV)
- Problem Setting -

■ Objective functions

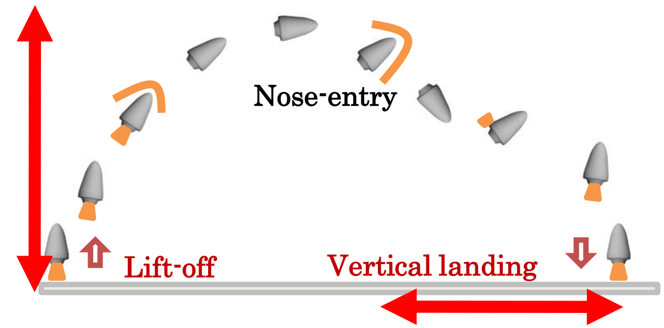
Ascent phase

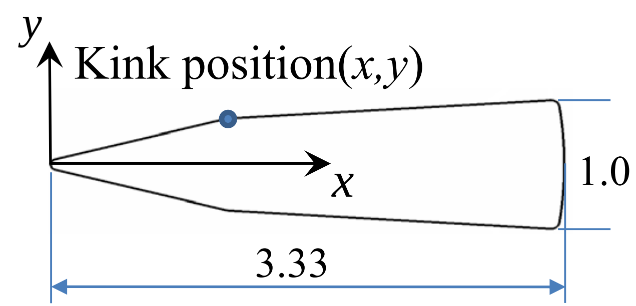
1. Drag to be minimized (M=2.0, AOA = 0)

Return phase

2. Maximum L/D to be maximized at subsonic condition (M=0.8)
3. Maximum L/D to be maximized at supersonic condition (M=2.0)

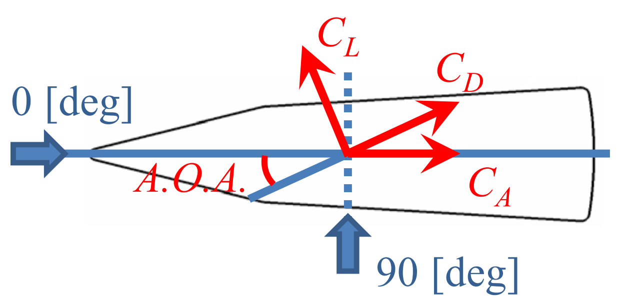
4. Body volume to be maximized





■ Design paramters

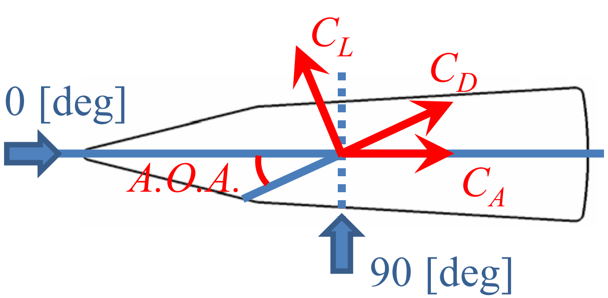
Kink position (x, y)



Extraction of design information using Proposed MOGP

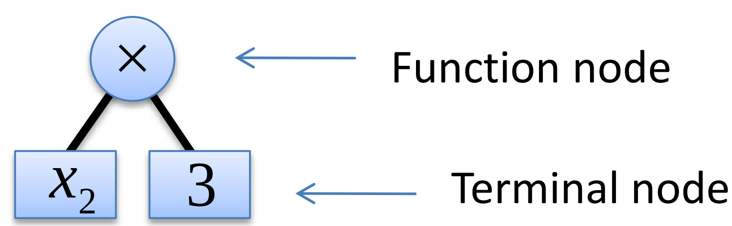
Objective functions

1. Maximization of squared correlation between $C_{A,\ M=2.0,\ AOA=0}$
2. Maximization of squared correlation between $L/D_{MAX,\ M=0.8}$
3. Maximization of squared correlation between $L/D_{MAX,\ M=2.0}$
4. Maximization of squared correlation between *Volume*
5. Minimization of the number of nodes in syntax tree

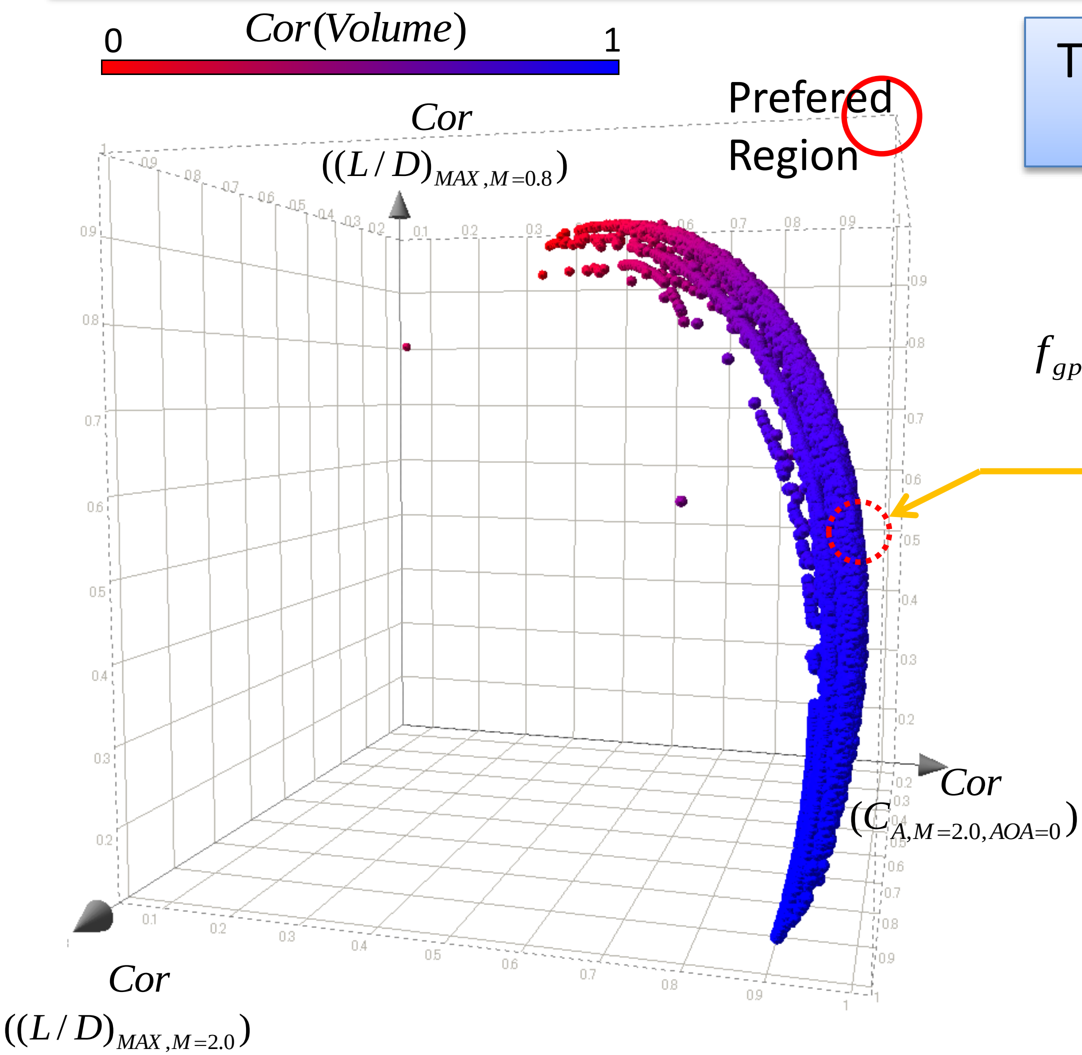


Computational conditions

Gene type	Tree expression
Generation	1500
Population	1500
Crossover ratio	0.8
Mutation ratio	0.1
Function sets	+ , - , * , /
Terminal sets	Design parameter (*x*, *y*), Constants [-1, 1]
Constraint	The number of nodes>1



Results of MOGP



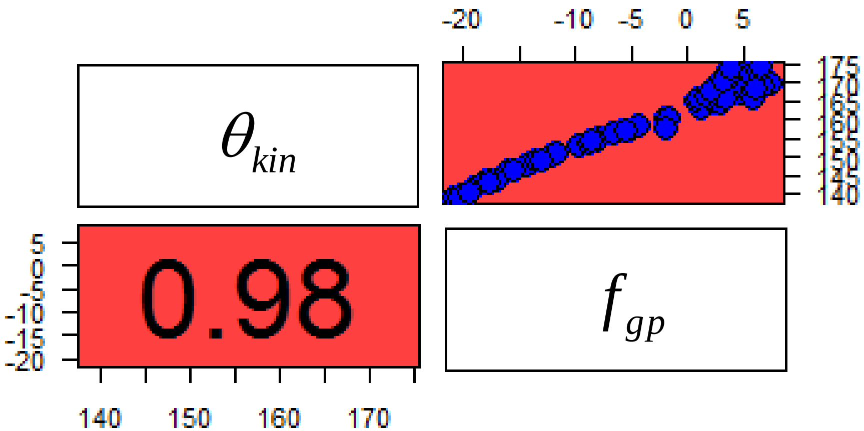
The maximum sum of all squared correlation

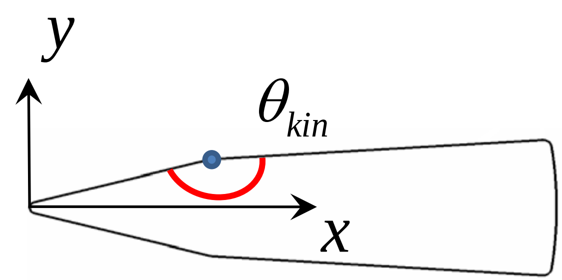
$$f_{gp} = -5.33y + 4.3xy + 4.3x$$
$$-1.28x^2 + xy^2 - 0.8x^2y - 0.33y^2$$

	Correlation
$C_{A,M=20,AOA=0}$	-0.99
$(L/D)_{MAX,M=0.8}$	0.44
$(L/D)_{MAX,M=2.0}$	0.98
$Volume$	-0.93

Results of MOGP

$$f_{gp} = -5.33y + 4.3xy + 4.3x - 1.28x^2 + xy^2 - 0.8x^2y - 0.33y^2$$





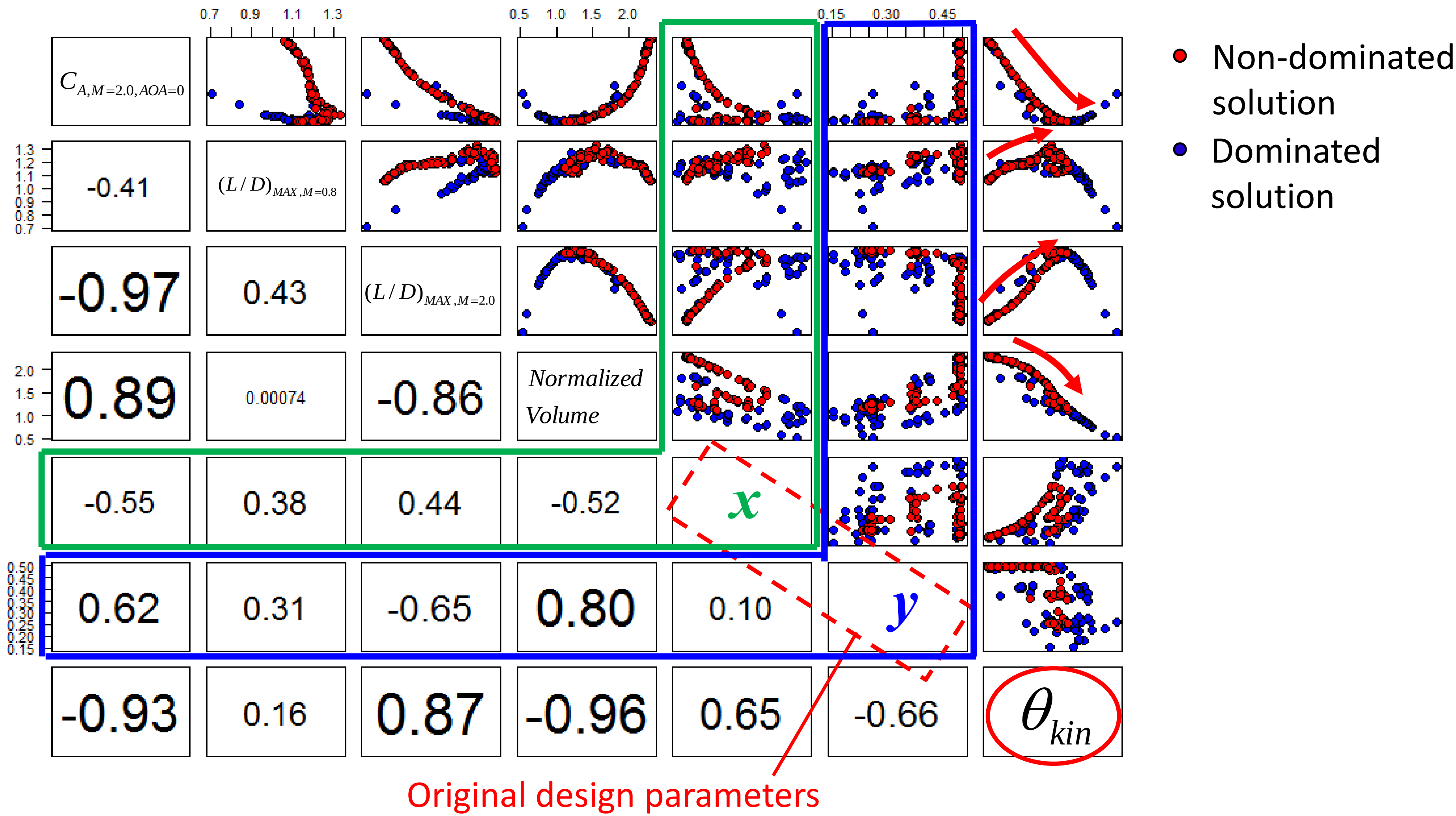
$$\theta_{kin} = 180 - \frac{180}{\pi}\arctan\left(\frac{y}{x}\right) + \frac{180}{\pi}\arctan\left(\frac{0.5-y}{3.33-x}\right)$$

$$= (188.53 - 16.8y - 0.74y^2 + O[y^3])$$

$$+ \frac{180}{\pi y} + 2.52 - 4.83y - 0.64y^2 + O[y^3])x$$

$$+ \left(-\frac{60}{\pi y^3} + O[y^2]\right) + O[x^4] + (-1)^{Floor(\ldots)}(-90 + O[y^3])$$

Results of MOGP



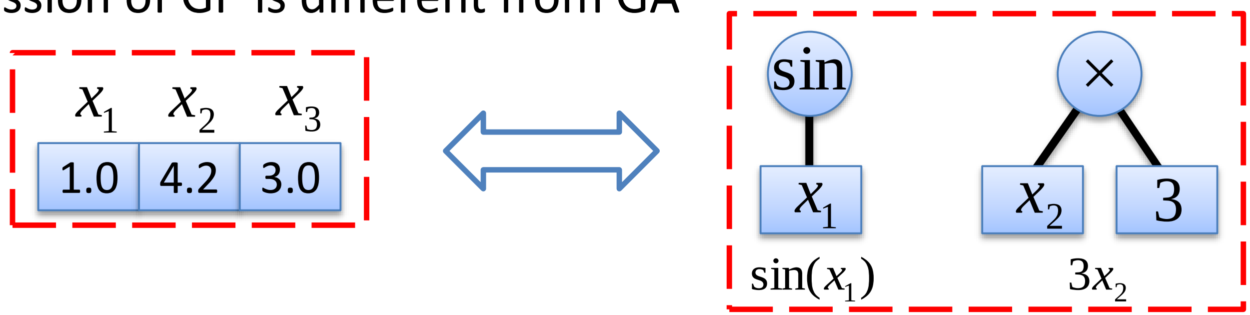
Conclusions

- A new multi-objective genetic programming for design exploration is proposed.
  - The unique feature of MOGP is the simultaneous symbolic regression to multiple variables using correlation coefficients.

- MOGP is applied to non-dominated solutions of the design optimization problem of a bi-conical shape reusable launch vehicle.
  - The result of proposed MOGP presents symbolic equations which have high correlation to zero-lift drag at supersonic condition, maximum lift-to-drag at supersonic condition and volume of shape at the same time.
  - These results also have high correlation to kink angle of the body geometry. It implies that proposed GP is capable of finding composite new design parameters.

About Genetic Programming

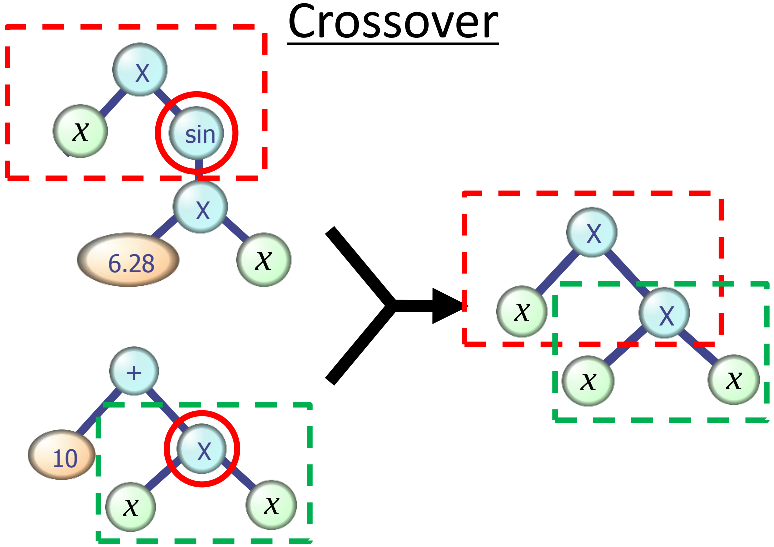
- Extension of Genetic Algorithm(GA) (*Koza, 1990*)
- Genome expression of GP is different from GA



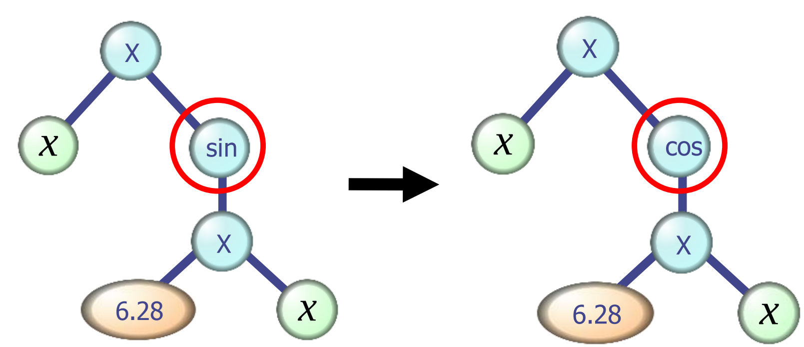
GA (Array) GP(Tree expression)

- Genetic operators

Crossover

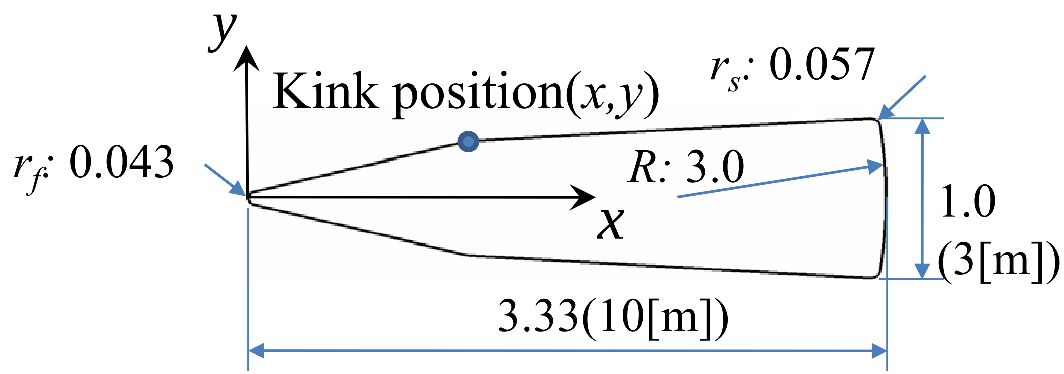


Mutation

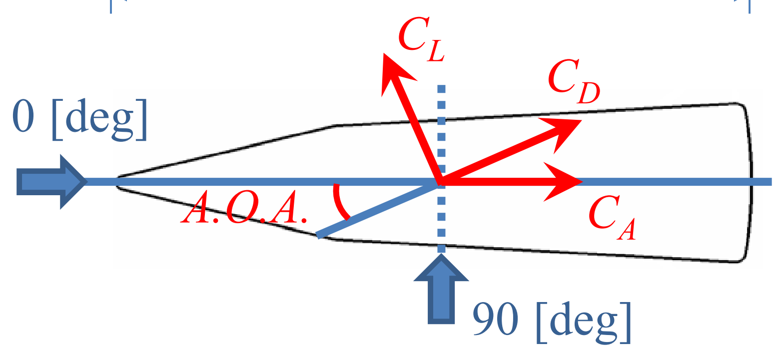


MODE for RVT
- Problem Definition -

■ Design parameters
- Kink position(x, y)



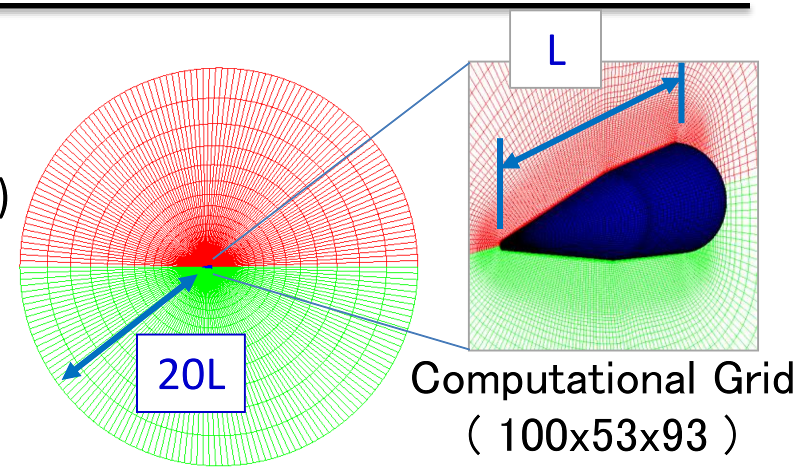
■ Constraint conditions
- Base diameter = 3[m]
- Length of the body = 10[m]
- Nose radius(r_f) = 0.043[m]
- Base corner radius(r_s) = 0.057[m]
- Base radius(R) = 3.0[m]



Computational methods and conditions

Computational Methods

- CFD
  - 3rd order MUSCL+SLAU (for shock instability)
  - Baldwin-Lomax
- MOEA
  - Non-dominated Solution GA-II (NSGA-II)

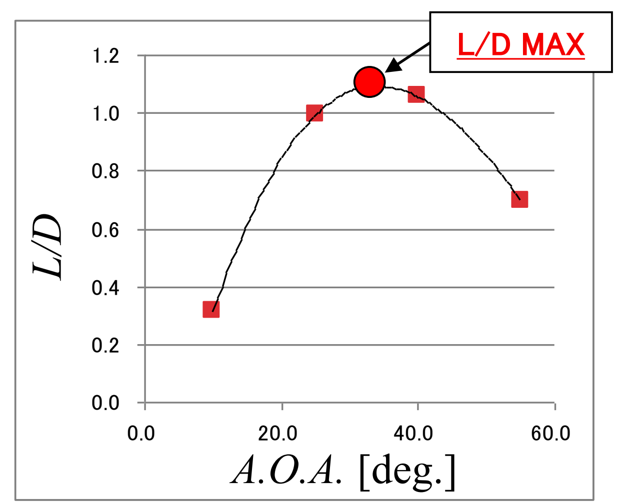


Computational Grid
(100x53x93)

Computational Conditions

- Flow analysis

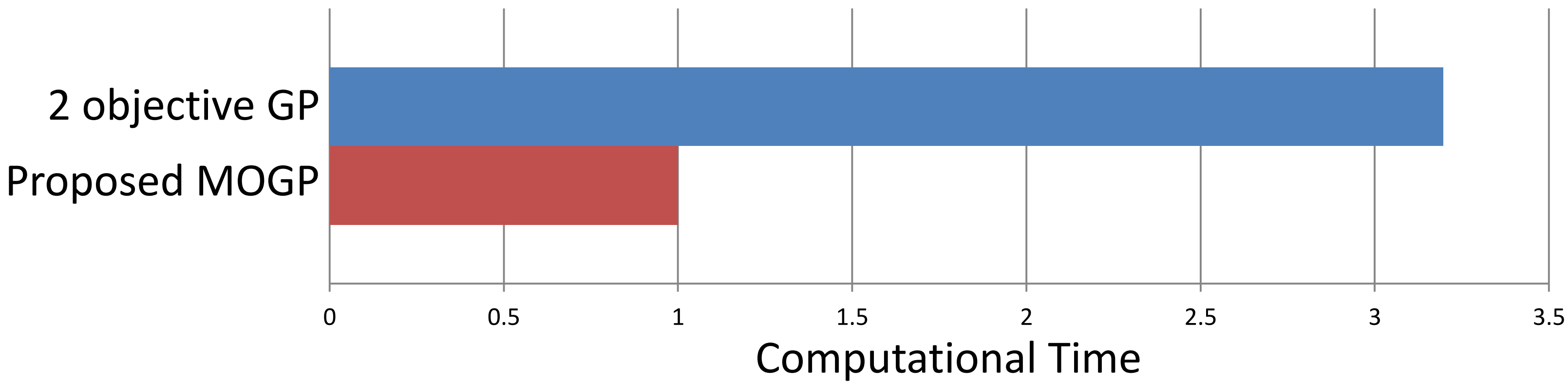
(Re = 1.0×10^7)	Mach number	Angle of Attack [deg.]
Ascent phase	2.0	0
Return phase	2.0	10,25,40
	0.8	10,25,40,55



- Optimization
  - Population size: 20
  - Generation size: 20

⇒ 400 body shapes(3200 CFD runs) are evaluated

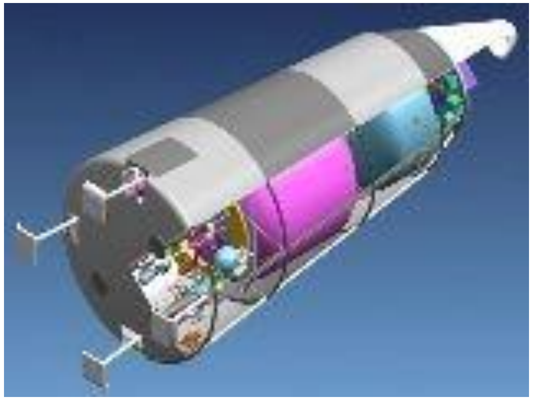
Comparison of computational time



Vertical Landing Reusable Launch Vehicle (RLV)

- One of the future space transportation systems
- Motivation of development
  - Low cost, reusability and reliability

Reusable Vehicle Testing(RVT)



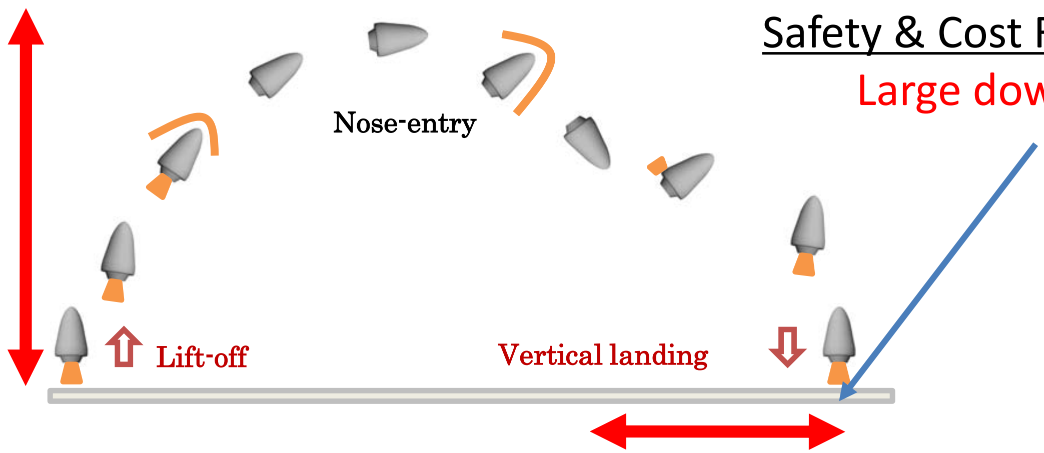
Reusable Sounding Rocket Vehicle(RSRV)

Vertical Landing Reusable Launch Vehicle (RLV)

■ Flight sequence & Design requirements

Mission Requirement

High ascent altitude



Safety & Cost Requirement

Large downrange

■ Important feature

- RLV is a vertical landing SSTO rocket.
- RLV flies over a wide range of the flight speed and attack angles.

It is important to understand relationship between shape parameter and flow field

More knowledge on aerodynamic shape design is necessary

A new type of Genetic Programming

1. Change the measure of accuracy

GP	Proposed GP
Minimization of mean absolute error $\min\left(\frac{\sum_i \lvert y_i - \hat{y}_i \rvert}{N}\right)$	**Maximization of correlation coefficient** $\max\left(\frac{\sum_i (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_i (y_i - \bar{y})^2}\sqrt{\sum_i (\hat{y}_i - \bar{\hat{y}})^2}}\right)$

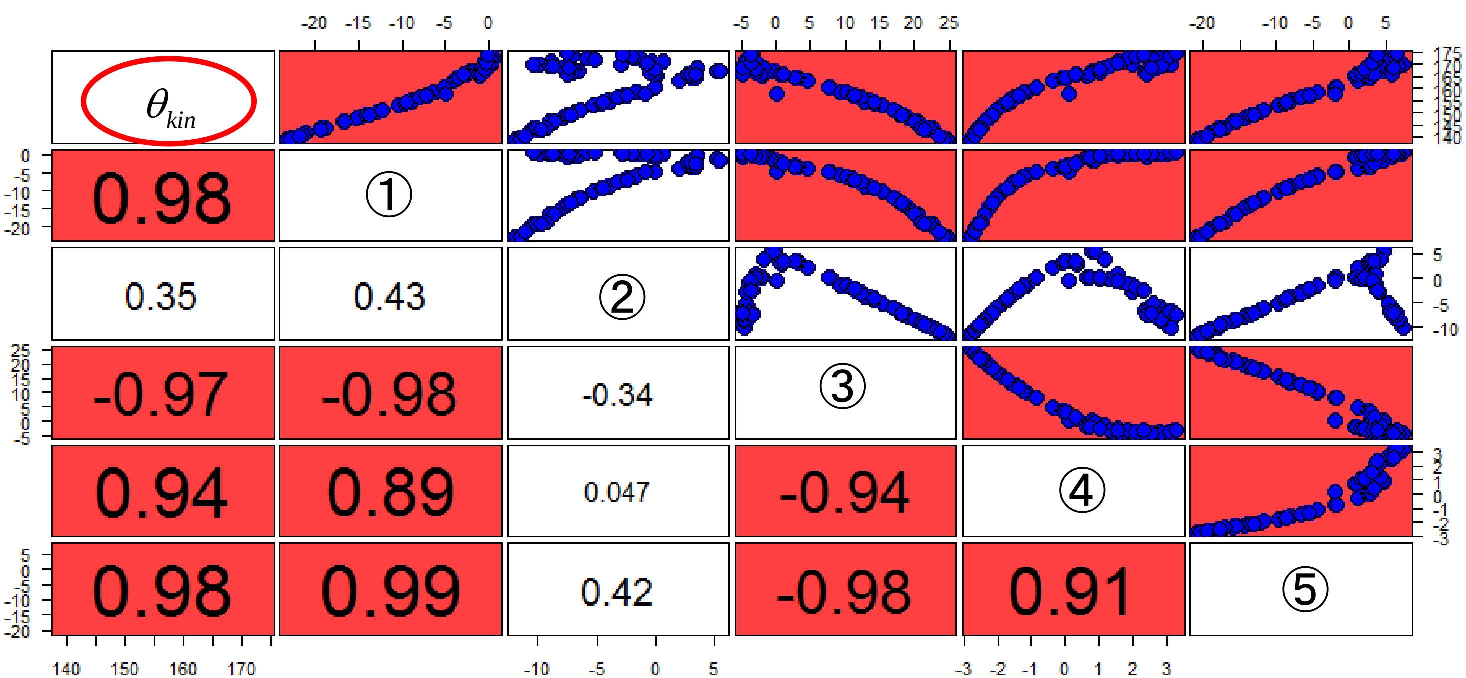
2. Change the number of objective functions

GP	Proposed GP
1) Minimization of mean absolute error between *obj*$_i$ **2) Minimization of the number of nodes**	**1) Maximization of correlation coefficient between *obj*$_1$** **2) Maximization of correlation coefficient between *obj*$_2$** ⋮ **N) Maximization of correlation coefficient between *obj*$_N$** **N+1) Minimization of the number of nodes**

***obj*$_i$** : *ith* objective function of MOO

Results of proposed MOGP

		x	y	x^2	y^2	xy	xy^2	x^2y	x^3	*Const.*
①	Maximum $Cor(C_{A,M=2.0,AOA=0})$	2.75	-4.1	-2.6	-2	3.1		-1	1	-1.65
②	Maximum $Cor(L/D_{MAX,M=0.8})$	0.7	0.6	-1	-4	4	2			0.4
③	Maximum $Cor(L/D_{MAX,M=2.0})$	-3.87	8.75	1.3	5	-0.1				-0.3
④	Maximum $Cor(Volume)$	1	-2		0.2	0.1				0.3
⑤	Maximum sum of all correlation	4.3	-5.33	-1.28	-0.33	4.3	-0.8			



Multi-Objective Design Exploration

- Problem Setting -

Problem Setting

↓

Multi-objective optimization

↓

Data mining methods

↓

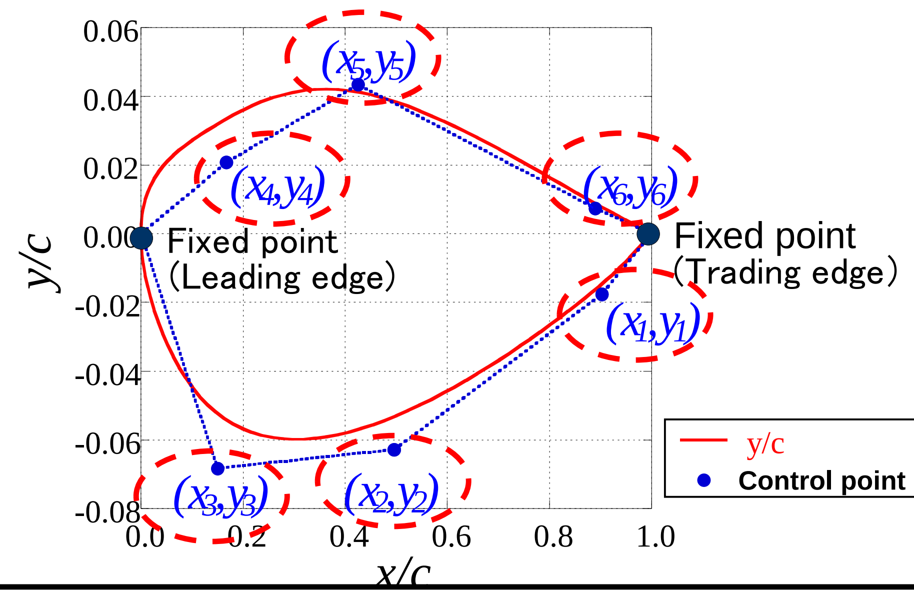
Extraction of design knowledge

e.g.) MODE for 2D wing shape

■ Objective functions

1. Maximization of lift coefficient(C_L)
2. Minimization of drag coefficient(C_D)

■ Design parameters



Multi-Objective Design Exploration

- Multi-objective optimization -

Problem Setting

↓

Multi-objective optimization

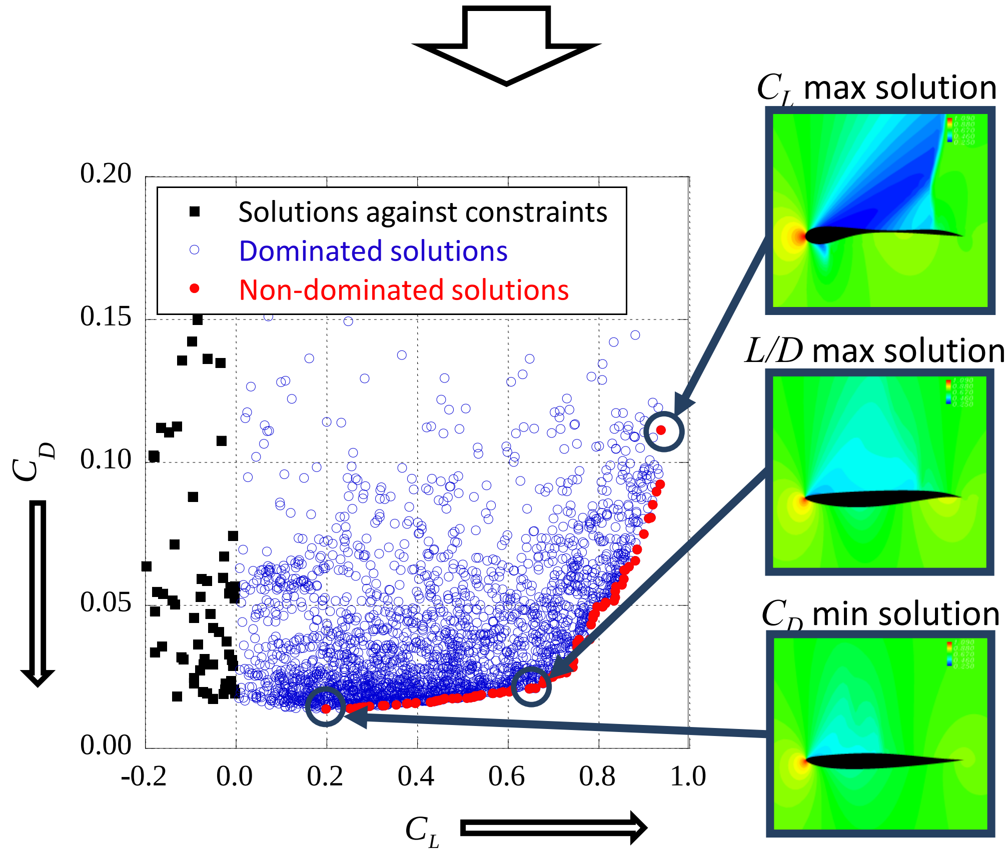
↓

Data mining methods

↓

Extraction of design knowledge

Non-dominated solutions are obtained by using multi-objective evolutionary algorithms

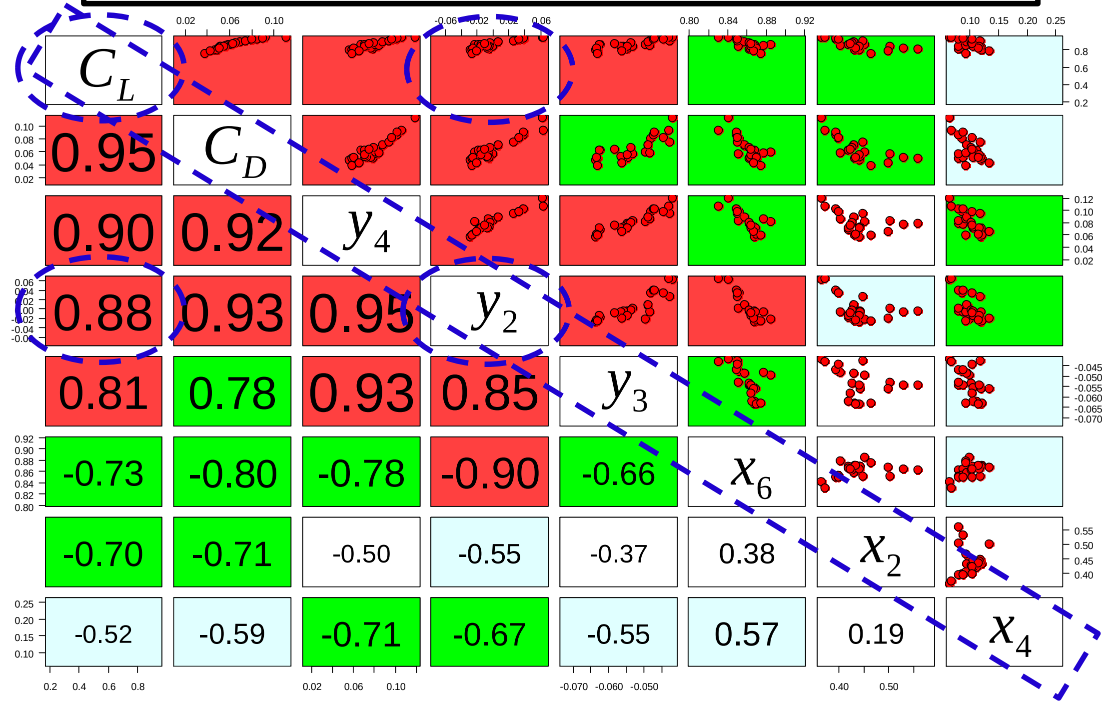


Multi-Objective Design Exploration

- Data mining methods -

- Problem Setting
- Multi-objective optimization
- Data mining methods
- Extraction of design knowledge

Scatter plot matrix with correlation coefficient



Multiple-regression analysis

$$C_L \approx 1.5y_4 - 0.67y_2 - 0.23y_3 - 0.33x_6 - 0.33x_2 + 0.21x_4$$

$$C_D \approx 1.5y_4 - 0.18y_2 - 0.57y_3 - 0.19x_6 - 0.24x_2 + 0.18x_4$$

Computational methods and conditions

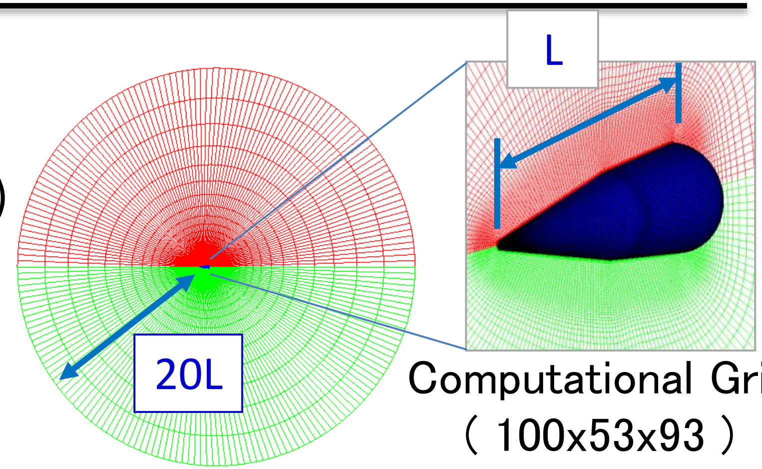
■ Computational Methods

■ CFD

- 3^{rd} order MUSCL+SLAU (for shock instability)
- Baldwin-Lomax

■ MOEA

- Non-dominated Solution GA-II (NSGA-II)

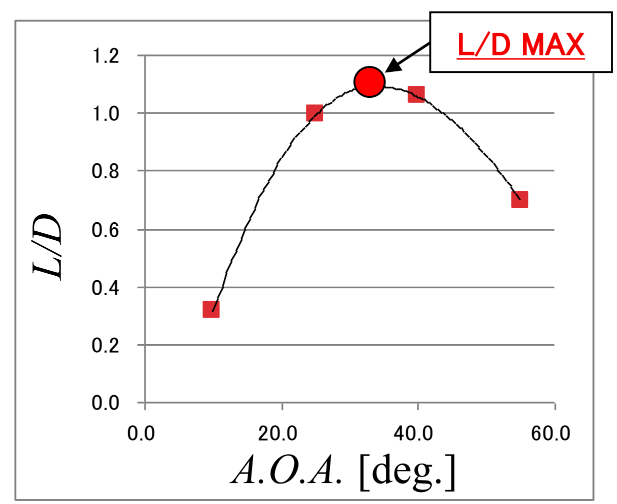


Computational Grid
(100x53x93)

■ Computational Conditions

■ Flow analysis

(Re = 1.0×10^7)	Mach number	Angle of Attack [deg.]
Ascent phase	2.0	0
Return phase	2.0	10,25,40
	0.8	10,25,40,55



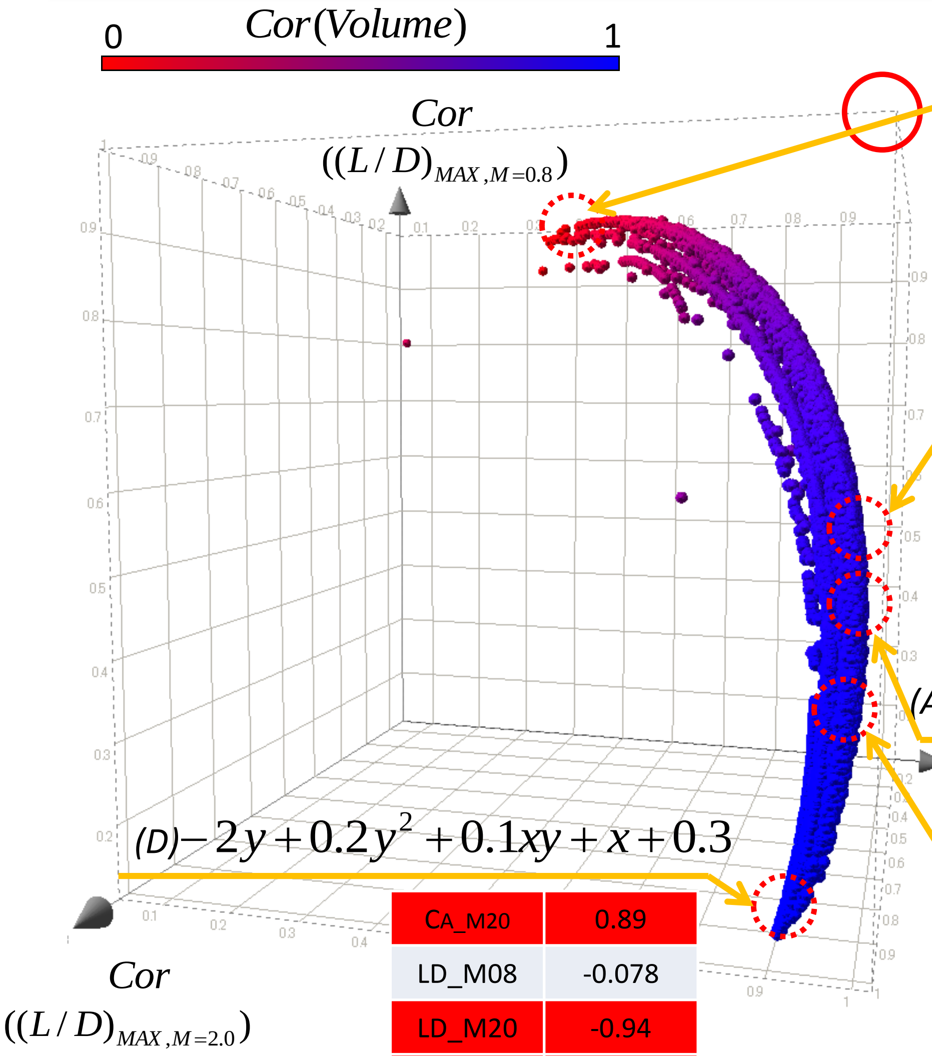
■ Optimization

- Population size: 20
- Generation size: 20

➡ 400 body shapes(3200 CFD runs) are evaluated

Results of MOGP

CA_M20	-0.44
LD_M08	0.97
LD_M20	0.35
Volume	-0.11



(B) $-4y^2+4xy+2xy^2-x^2+0.7x+0.6y+0.4$

(E) The maximum sum value of all squared correlation

$$-5.33y+4.3xy+4.3x-1.28x^2+xy^2-0.8x^2y-0.33y^2$$

CA_M20	-0.99
LD_M08	0.44
LD_M20	0.98
Volume	-0.93

CA_M20	-1.0
LD_M08	0.46
LD_M20	0.98
Volume	-0.91

(A) $-4.1y+3.1xy+2.75x-2.6x^2-2y^2-x^2y+x^3-1.65$

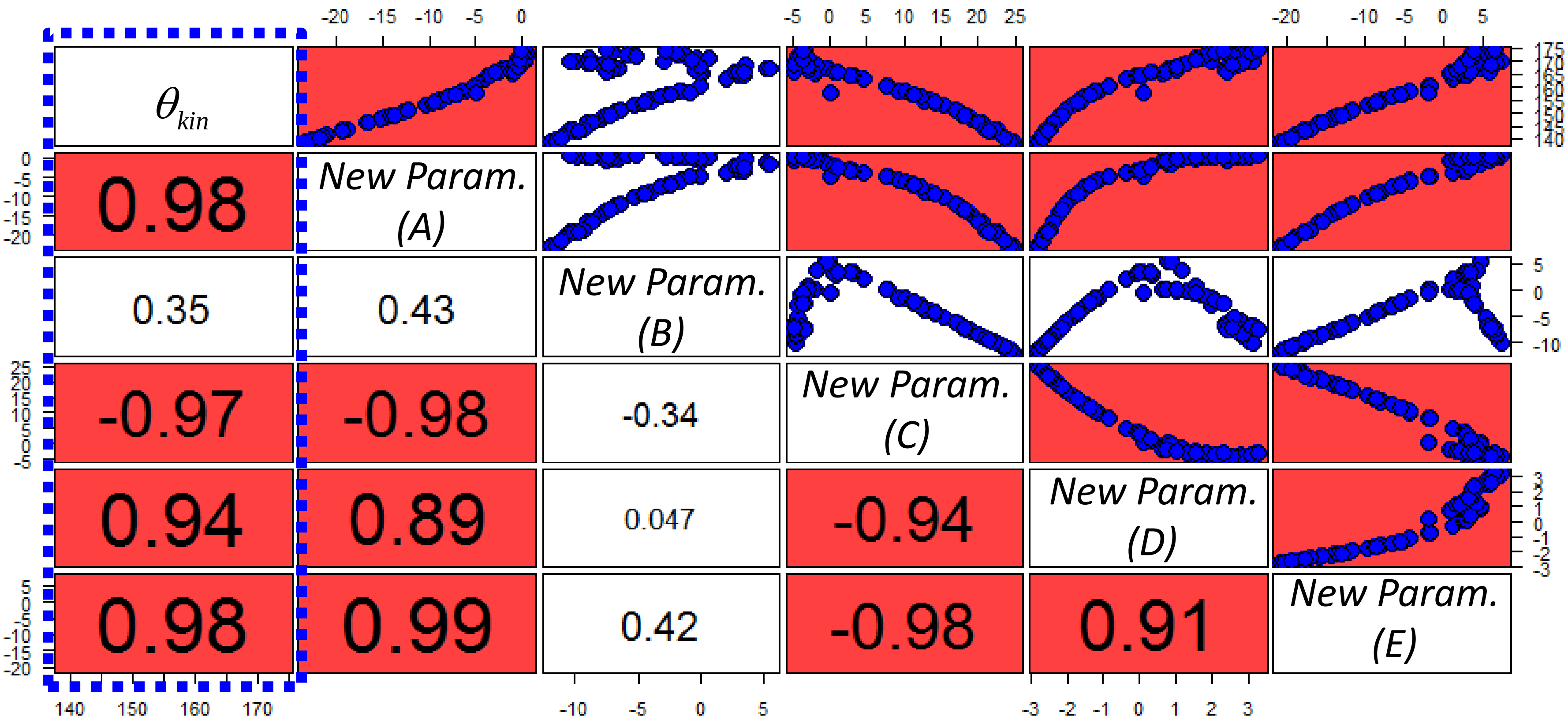
CA_M20	0.98
LD_M08-	-0.36
LD_M20	-1.0
Volume	0.96

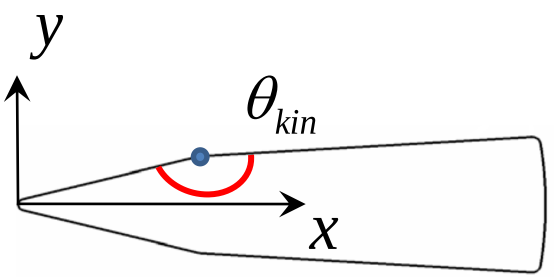
(C) $8.75y+5y^2-5xy-3.87x+1.3x^2-0.4$

(D) $-2y+0.2y^2+0.1xy+x+0.3$

CA_M20	0.89
LD_M08	-0.078
LD_M20	-0.94
Volume	1.00

Results of MOGP



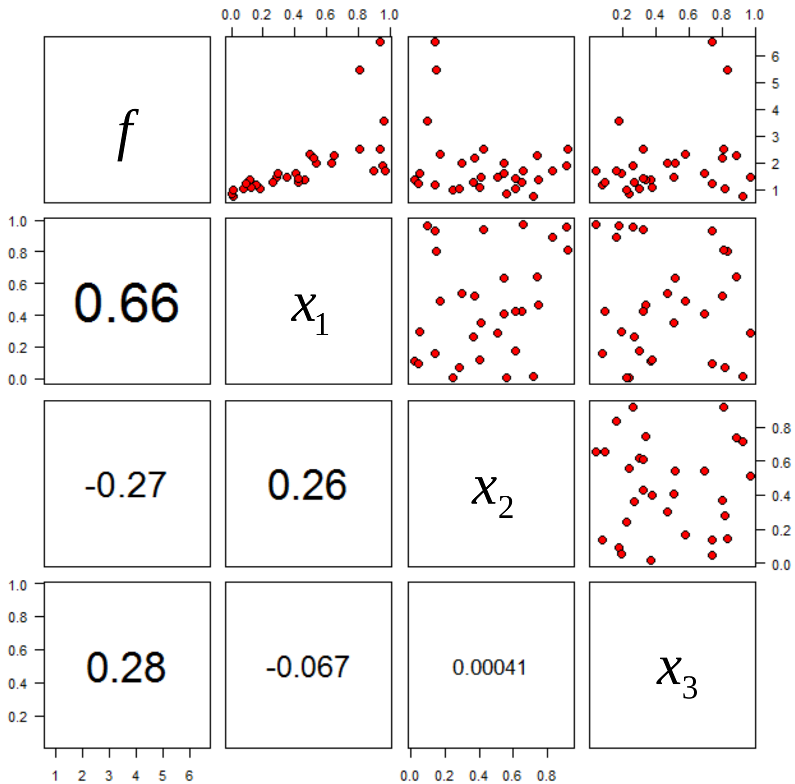


$$\theta_{kin} = 180 - \frac{180}{\pi}\arctan\left(\frac{y}{x}\right) + \frac{180}{\pi}\arctan\left(\frac{0.5-y}{3.33-x}\right)$$

Example of GP symbolic regression

Test function : $f(x_1, x_2, x_3) = \sin(x_1) + \cos(x_2) + \frac{x_1^{\;2} x_3}{x_2} + x_1 x_2 x_3$

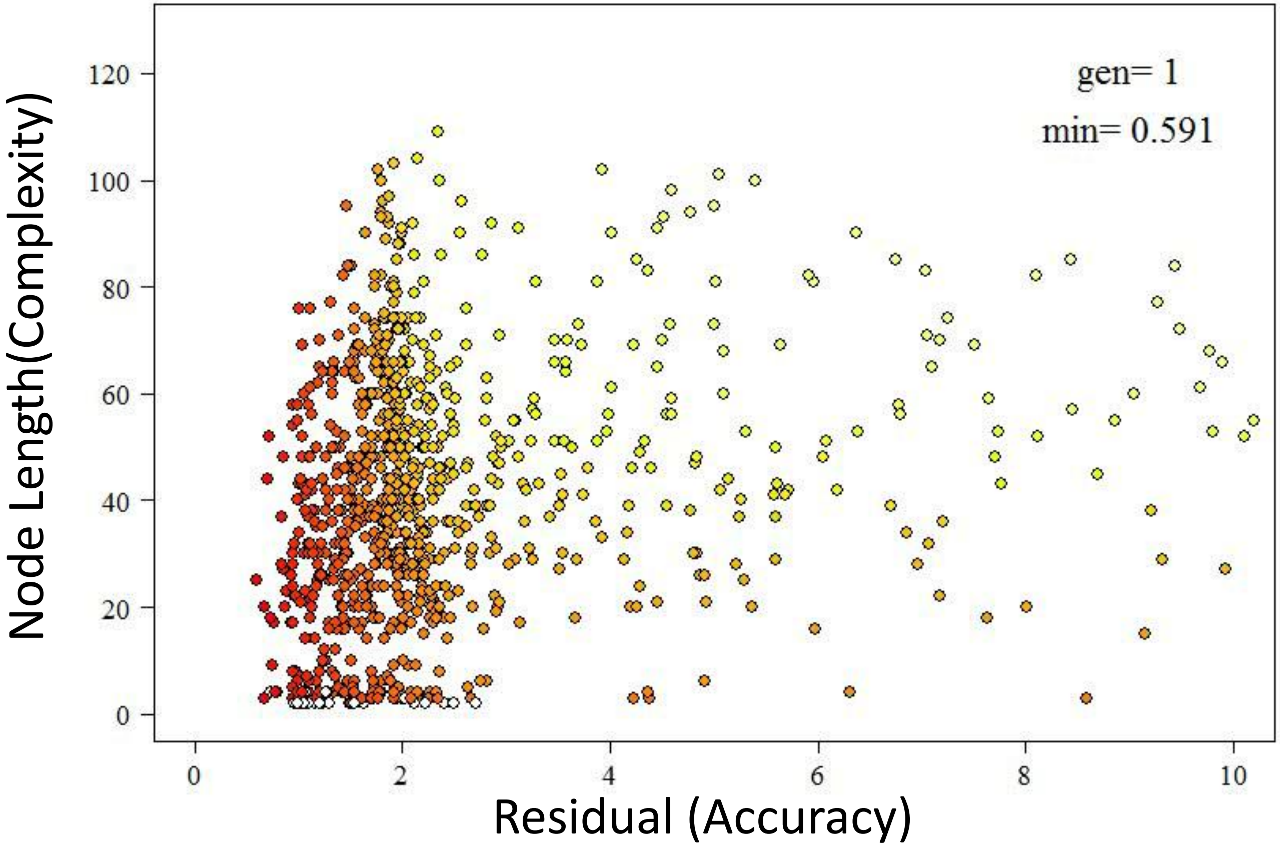
Data set



Computational Condition

Gene type	Tree expression
Generation	1000
Population	1000
Crossover ratio	0.8
Mutation ratio	0.2
Function sets	+, - , *, /, sin, cos
Terminal sets	Design parameters, Constants [-1, 1]
Constraint	The number of nodes>1

Example of GP
- Result -



574th generation → $f(x_1, x_2, x_3) = \sin(x_1) + \cos(x_2) + \frac{x_1^{\ 2} x_3}{x_2} + x_1 x_2 x_3$

symbolic regression example

- Test function

$$f_1(x_1, x_2, x_3) = 10x_1^{\ 3} + 5x_2x_3 + 10$$

$$f_2(x_1, x_2, x_3) = 10x_2^{\ 3} + 5x_3x_1 + 10$$

$$f_3(x_1, x_2, x_3) = 10x_3^{\ 3} + 5x_1x_2 + 10$$

$$x_1, x_2, x_3 = [-1,1]$$

- Data set
  - 40 sample points (random)

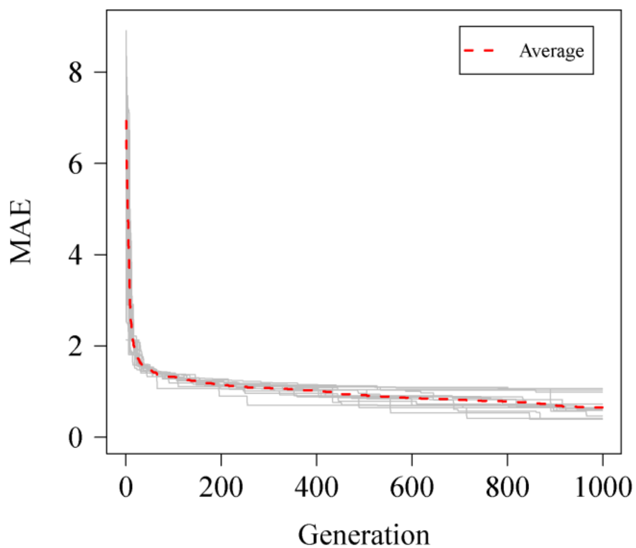
- GP
  - 3 type of GP
    - TOGP-MAE（2 Objective GP, Mean Absolute Error)
    - TOGP-SCC（2 Objective GP, Squared Correlation Coefficient)
    - MOGP-SCC（Multi Objective GP, Squared Correlation Coefficient)

  - 1000 Individuals, 1000 Generations
  - 15 Trial

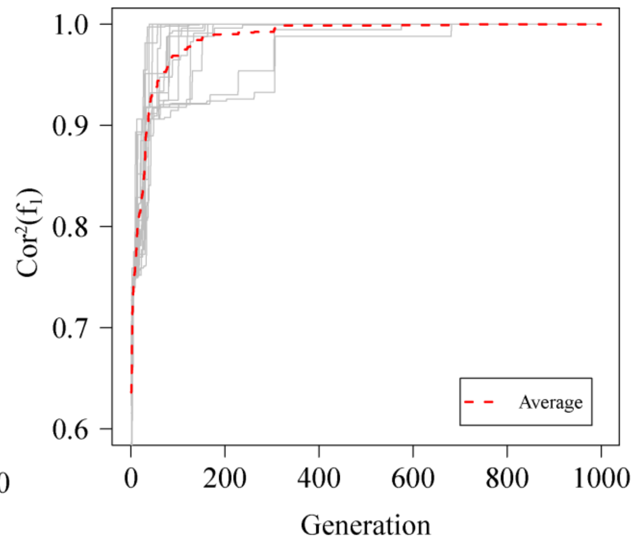
Results

- History of accuracy

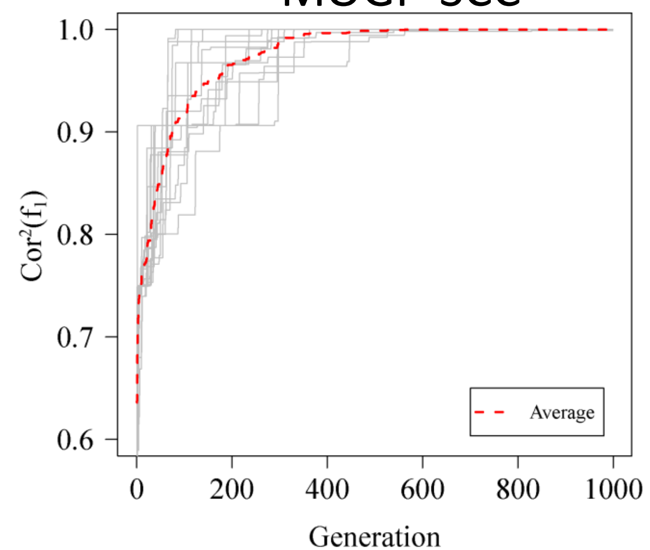
TOGP-MAE



TOGP-SCC



MOGP-SCC



- Comparison of computational time