



---

# データマイニングを用いたエンドミル切削条件決定の効果

---

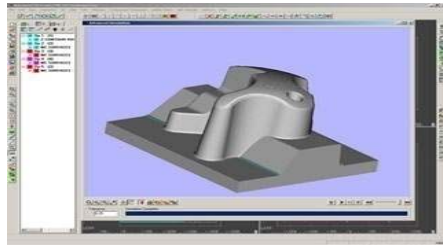
同志社大学大学院 工学研究科  
児玉 紘幸

---

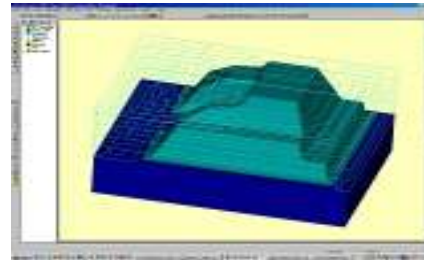
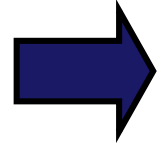
# 目次

- ◆ 背景
- ◆ データマイニングとは
  - カタログマイニングプロセス
  - 非階層・階層クラスタリングの併用効果
- ◆ 研究目的
- ◆ カタログマイニング
  - 源データの取得
  - 工具形状のモデル化
  - 形状を代表する変数の選択
  - 導出された切削条件式の妥当性の検討
  - 知識の解釈
- ◆ 結言

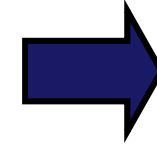
# 背景



CAD



CAM



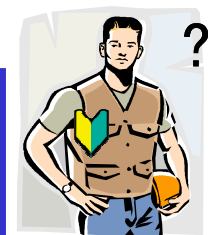
マシニングセンタ

近年のCAMシステムの発達により、熟練技術者でなくても容易にNCプログラムを生成することが可能となった

しかし・・・

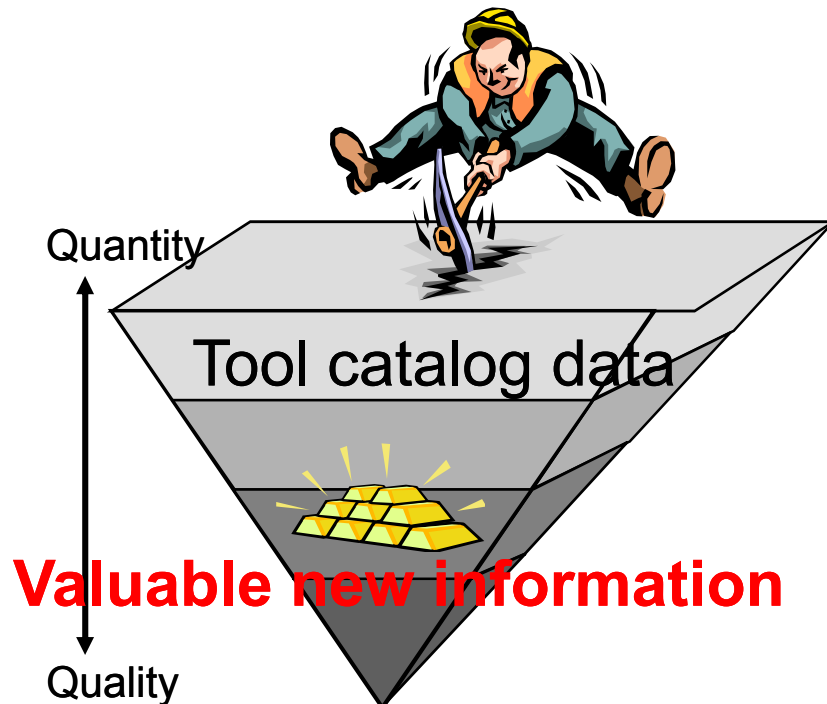
- ◆加工をする上で重要な使用工具や切削条件の決定を、熟練技術者の知識や経験に頼っているのが現状
- ◆熟練技能者の知識や経験は、暗黙知である場合が多く非熟練技能者には伝承されにくい

➤公開されている知識やデータを活用して、使用工具や切削条件の決定を支援するシステムを構築していくことが必要であると考えられる



# 目的

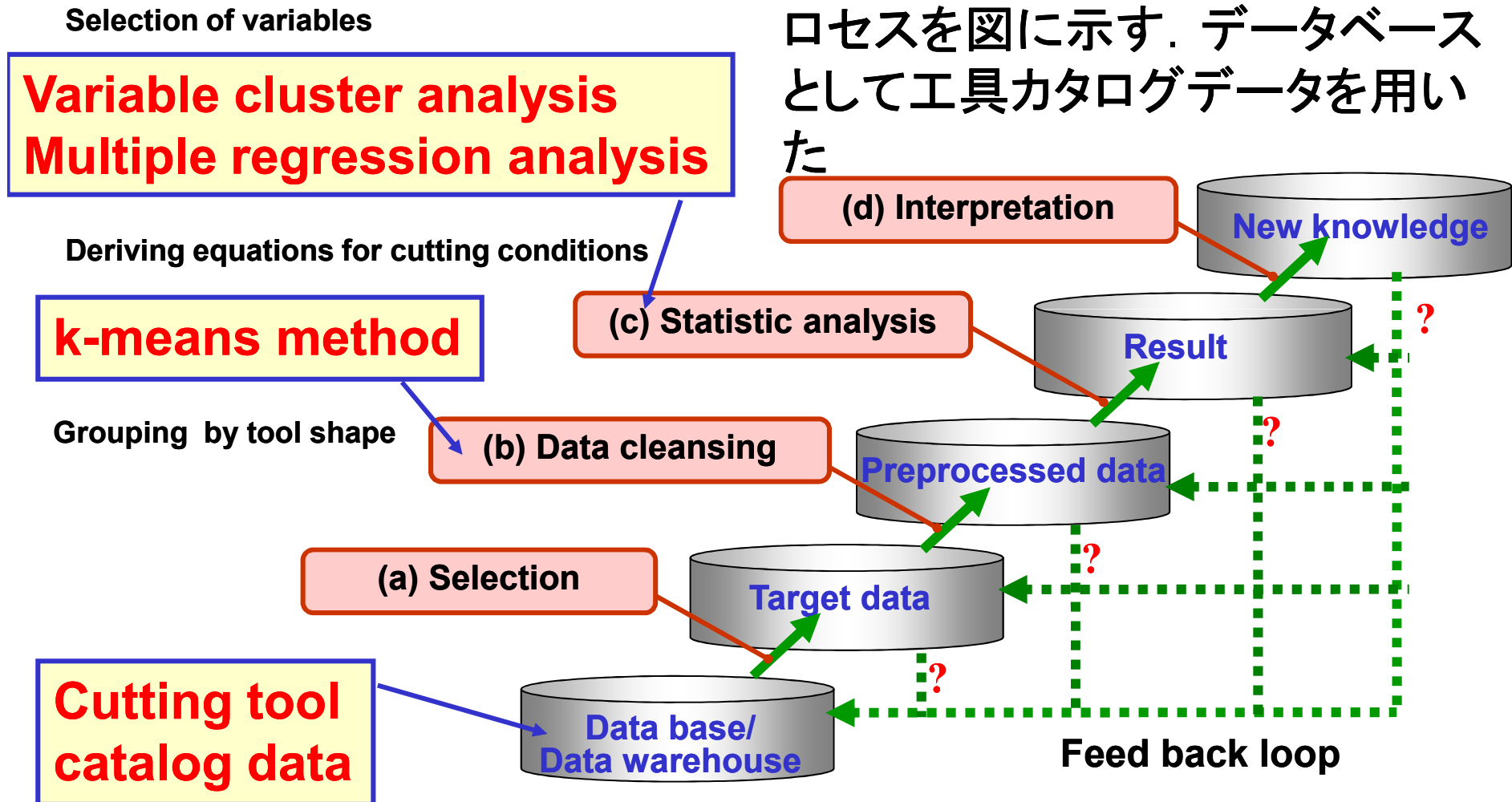
- 工具カタログは工具に対する膨大な知識を集めた加工技術の宝庫であるため、良質なデータベースであるといえる
- 本研究では膨大なデータを有する工具カタログに対して、データマイニング手法を適用した切削条件決定支援システムを提案している





# カタログマイニングプロセス

➤ 提案するカタログマイニングプロセスを図に示す. データベースとして工具カタログデータを用いた



## カタログマイニングプロセス

- データマイニングにおいてとくに重要なのはデータの獲得, 選択, データクレンジングである.
- データをデータマイニングに使える形に変換する部分が新知識発見のために必要な全プロセスに占める割合は70~80%であるといわれている
- しかし工具カタログデータでは, すでに工具メーカーで実験を繰り返した結果に基づく良質なデータが大量に存在していると考えられるため, 解析者にとってデータの獲得, 選択に時間を割く必要はない
- またカタログには, すでに数値で表示されている工具データが多くデータクレンジングの多くを省略できるものと考えられる. そのため, カタログマイニングではデータのクラスタリングがとくに重要になると考えられる

## カタログマイニングプロセス

- (株)数理システム製のVisual Mining Studioを用いて、データ構造解析手法であるk-means法によって工具形状によるビジュアル的なクラスタ分けを行った
- 次に統計解析手法として、階層型解析手法である変数クラスタ分析により、変数間の関係をビジュアル的に表現することでデータ構造の階層を把握し、重回帰分析により切削条件式を導出し、工具形状と切削条件について考察する
- カタログマイニングシステムは、他の製造分野に適用されるマイニング手法に比べて、簡単で早く、しかも理解しやすいため、データマイニングに精通していない技術者でも取り扱いやすいのが特徴である



## 非階層・階層クラスタリングの併用効果

➤ 工具カタログデータに対して、単に重回帰分析を行うだけでは埋もれていた知識を、階層型の変数クラスタ分析、および非階層型のK-means法を併用してビジュアル化することで、より容易に発掘できる

解析手法	利点	欠点
非階層型クラスタリング K-means法	<ul style="list-style-type: none"> <li>似た者同士をビジュアル的に分類できる</li> </ul>	<ul style="list-style-type: none"> <li>目的変数がない</li> <li>初期値によって結果が異なる</li> </ul>
階層型クラスタリング 変数クラスター分析	<ul style="list-style-type: none"> <li>相関の高い変数同士の構造をビジュアル的に判断できる</li> </ul>	<ul style="list-style-type: none"> <li>目的変数への寄与率はわからない</li> </ul>
多変量解析 重回帰分析	<ul style="list-style-type: none"> <li>目的変数の値を説明変数の式で表すことができる</li> <li>説明変数の目的変数への寄与率を知ることができる</li> </ul>	<ul style="list-style-type: none"> <li>相関の高い変数は同時に組み込めない</li> <li>データ数が多いと精度の良い回帰式が作成できないことがある</li> </ul>

## 研究目的

本研究では、工具カタログに記載されている超硬スクエアエンドミルをデータベースとしてカタログマイニングを行うことで、

- 専門家も気付かない新知識を発見すること.
- 非熟練技能者にも実用的な切削条件を導出できる予測式を導出すること.

以上の二点を目的とする.

# カタログマイニングプロセス

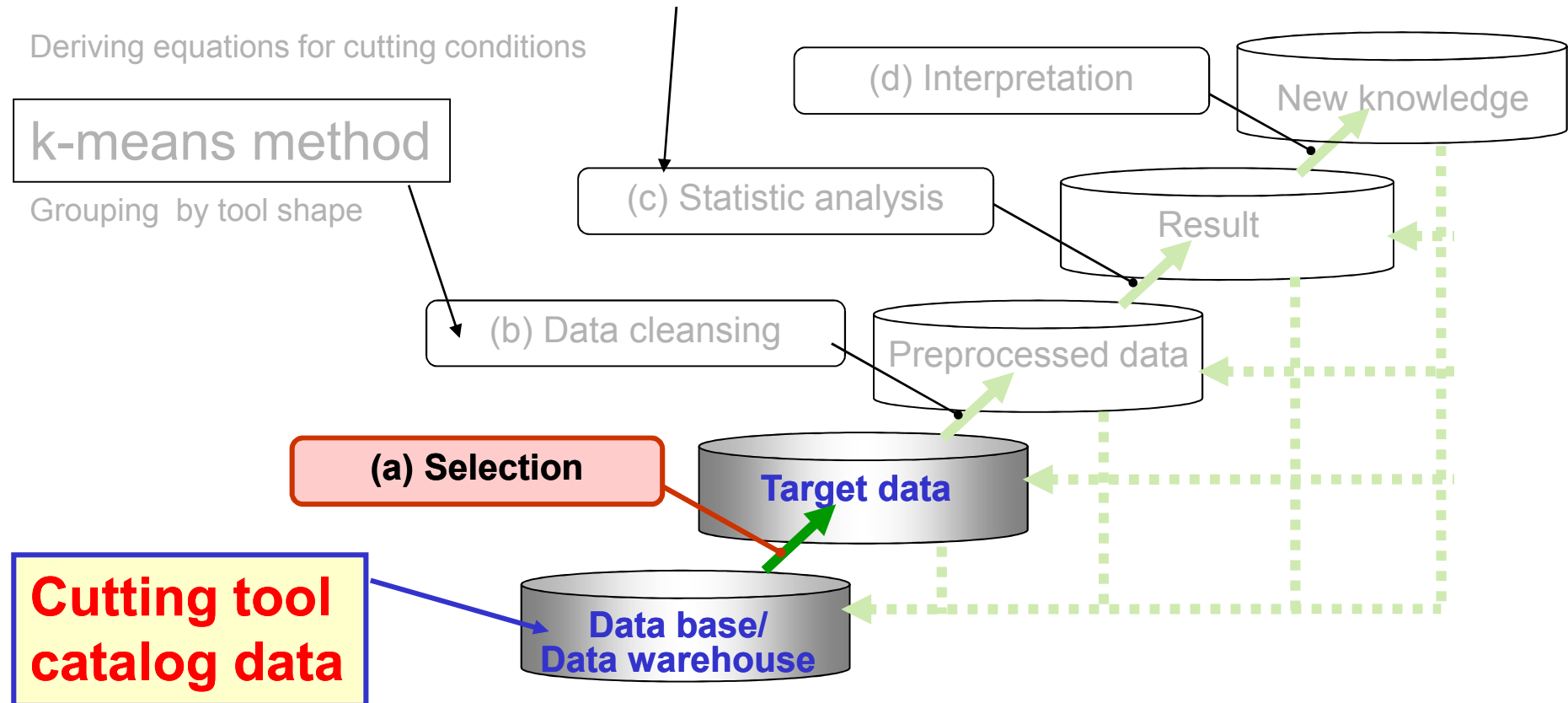
Selection of variables

Variable cluster analysis  
Multiple regression analysis

Deriving equations for cutting conditions

k-means method

Grouping by tool shape



# 原データの取得

## データベース

カタログ : 国内大手メーカーA社(2009-2010年度版)

選択した工具: スクエアエンドミル(データ総数2235項)

被削材: 炭素鋼, 合金鋼, 焼き入れ鋼(荒加工)HRC32~60

工具カタログに記載

### 説明変数

外径 $D$ , シャンク径 $D_s$ , 全長 $L$   
刃長 $l$ , 刃数 $z$ , ねじれ角 $\theta$   
被削材硬さ $HRC$ ,  
表示価格 $PRICE$

NCプログラムに入力する値

### 目的変数

切削速度 $V$   
一刃当りの送り量 $f$   
軸方向切り込み量 $A_d$   
径方向切り込み量 $R_d$

(寸法系の単位はmm, 角度はdegを用いる)

重回帰分析によって切削条件決定式を導出する

# カタログマイニングプロセス

Selection of variables

Variable cluster analysis  
Multiple regression analysis

Deriving equations for cutting conditions

**k-means method**

Grouping by tool shape

(b) Data cleansing

(a) Selection

Cutting tool catalog data

(c) Statistic analysis

(d) Interpretation

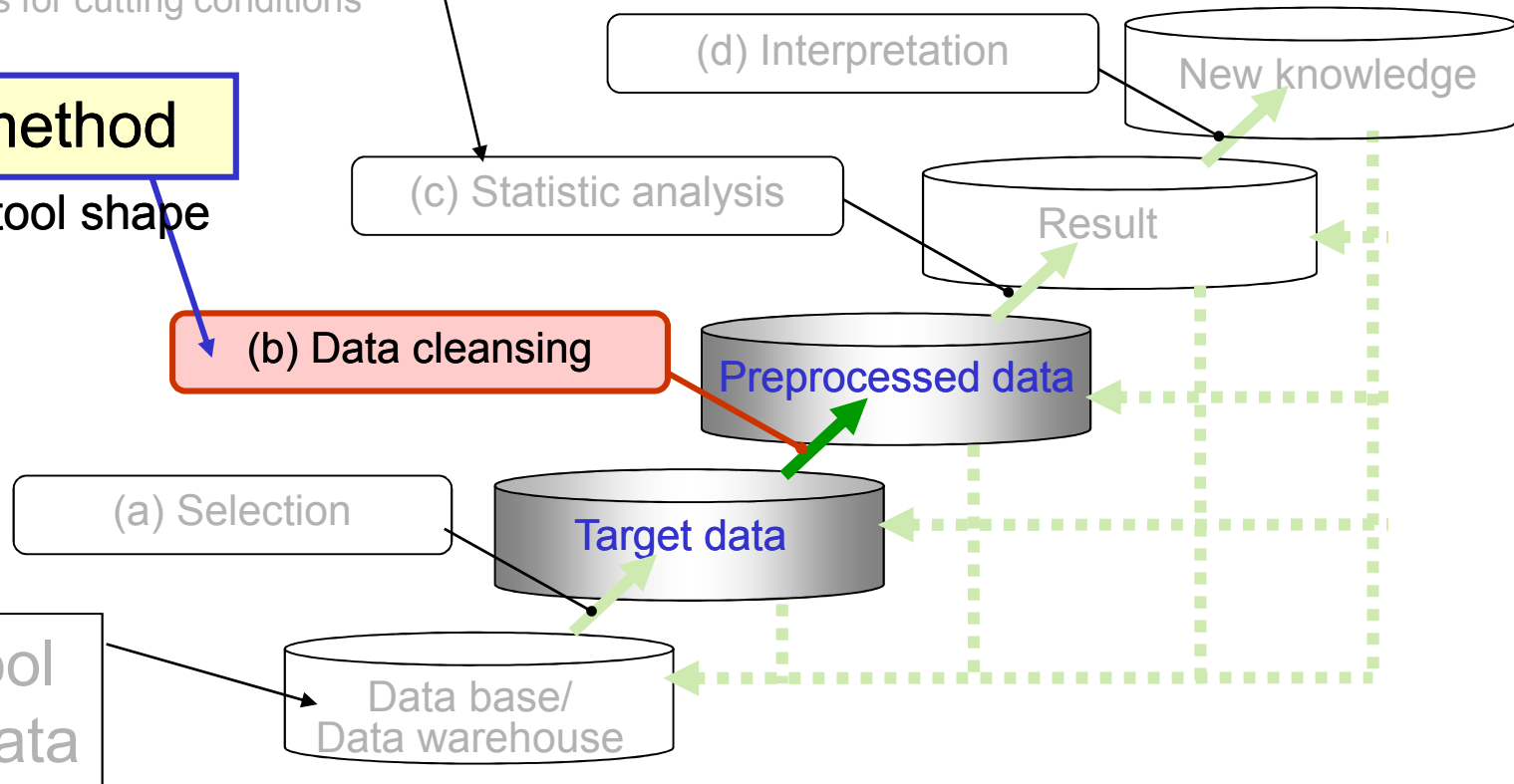
New knowledge

Result

Preprocessed data

Target data

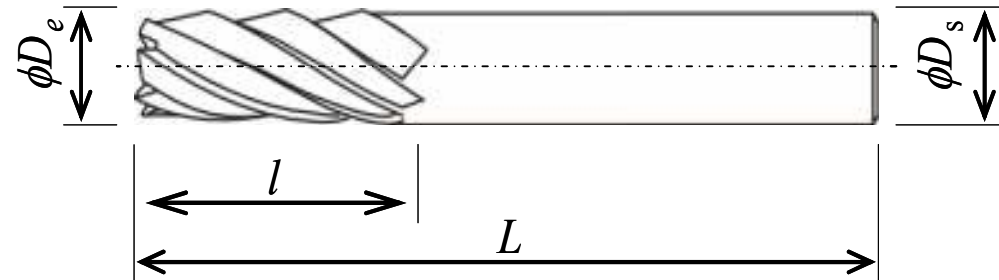
Data base/  
Data warehouse



# 工具形状のモデル化

## 非階層的手法: K-means法

Number of cutting tooth	Sectional area rate	De
2	0.55	0.742D
3	0.58	0.762D
4	0.62	0.787D



$L$ : 全長  $l$ : 刃長

$De$ : 相当径(相当質量換算)  $D_s$ : シャンク径

➤ 刃数を考慮して、刃部を同面積の円に置き換えた時の円の直径を相当径  $De$  [mm] と定義した。多数のエンドミルを準備し、その質量と刃長の実測値より  $De$  を調べた。その結果、各刃数における刃部の断面積および円の直径は表の関係であることがわかった。

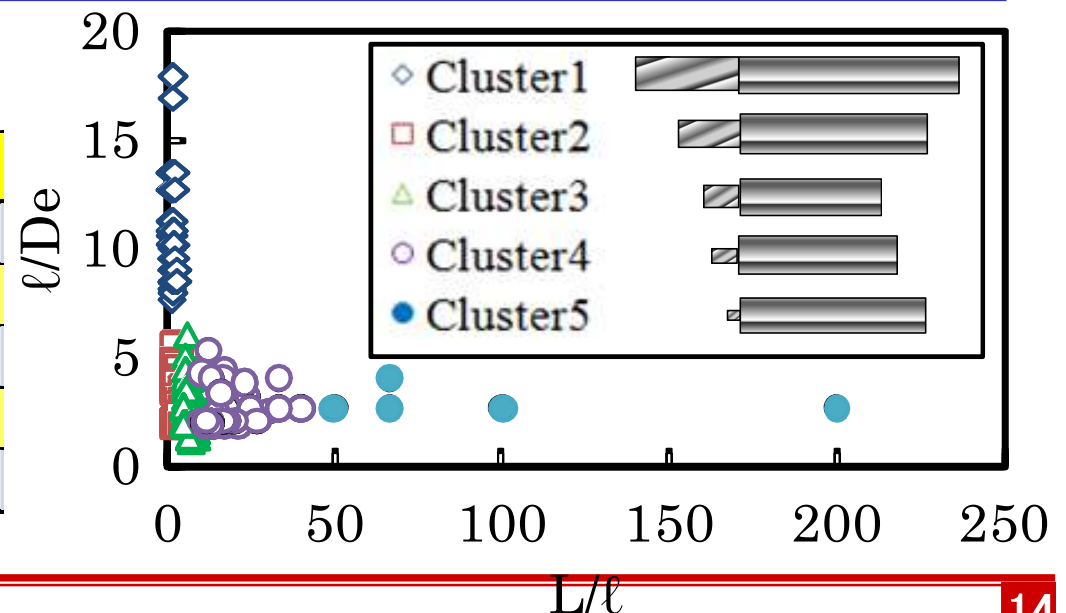
$L/l$ ,  $l/De$ ,  $D_s/De$ , 刃部コーティング の4つの係数を導入

工具形状から5つのクラスタに分類した

## 工具形状のモデル化

- 各クラスターのデータ数と $L/\ell$ ,  $\ell/D_e$ ,  $D_s/D_e$ の3つの形状変数の平均値を表に示す.
- クラスタ1は一般的形状, クラスタ5は先の細い形状であり, その間にクラスタ2, そして3, 4が属する. 実際に分類された各クラスターの $L/\ell$ と $\ell/D_e$ の関係を図に示す. ある工具がどのクラスターに属するかを調べるときは, これらの2変数の値に着目するだけでよいことがわかった

Cluster	Data size	$\ell/D_e$	$L/\ell$	$D_s/D_e$
1	152	2	11	1
2	1292	3	3	1
3	577	6	3	1
4	181	15	2	3
5	33	77	3	20



# カタログマイニングプロセス

Selection of variables

Variable cluster analysis  
Multiple regression analysis

Deriving equations for cutting conditions

k-means method

Grouping by tool shape

(c) Statistic analysis

(d) Interpretation

New knowledge

(b) Data cleansing

Preprocessed data

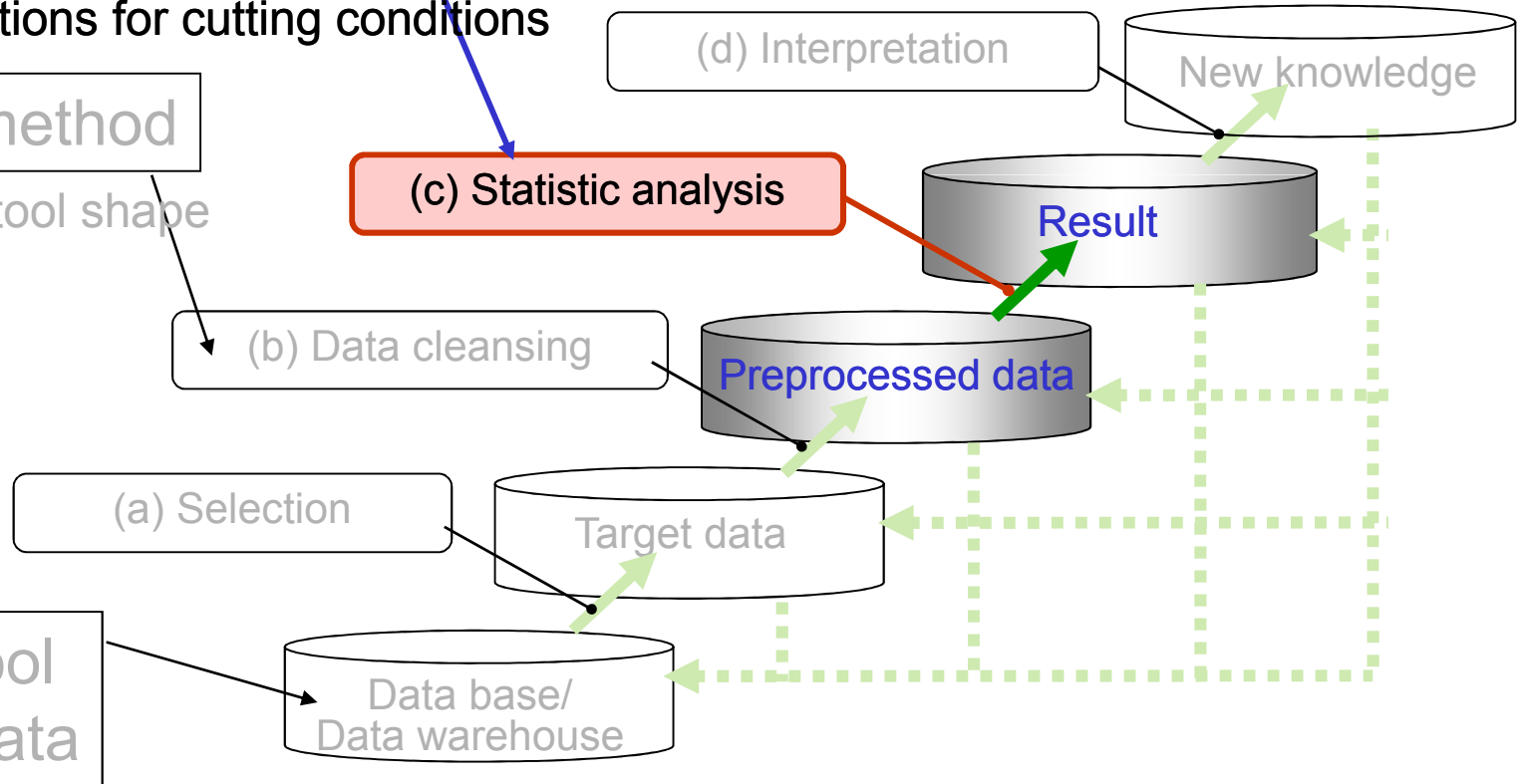
Result

(a) Selection

Target data

Cutting tool catalog data

Data base/  
Data warehouse

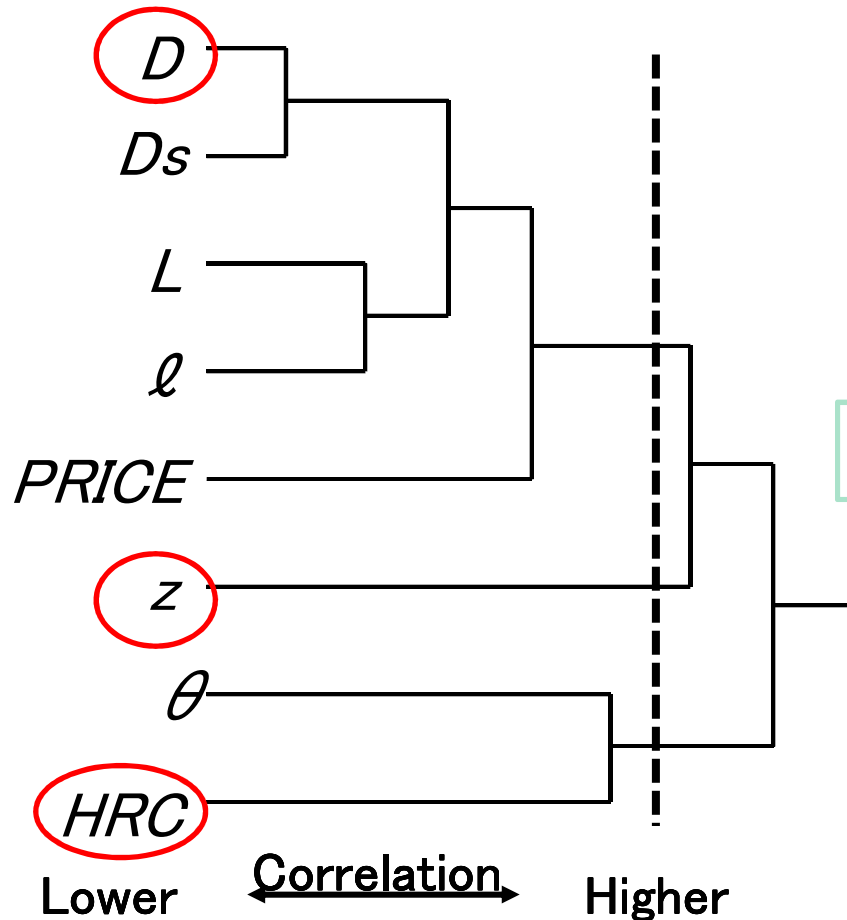




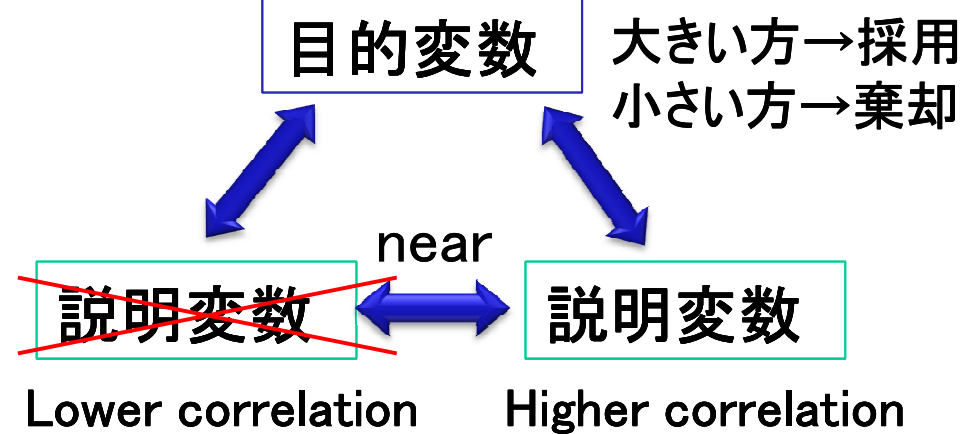
# 形状を代表する変数の選択

## 樹形図

例. Cluster2 側面加工



## 目的変数との相関係数



2変数間の距離  $d_{xy} = \sqrt{2(1 - r_{xy})}$

相関係数  $r_{xy} = \frac{\sum_{r=i}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{r=i}^n (x_i - \bar{x})^2} \sqrt{\sum_{r=i}^n (y_i - \bar{y})^2}}$

最終的に残った変数で  
重回帰分析を行う

## 形状を代表する変数の選択

➤ 樹形図をもとに相関の高い説明変数同士において、目的変数との相関係数を比較して相関の低い方を削減した。図中の点線で示されるカッティング線により、説明変数として刃数 $z$ と被削材硬さ  $HRC$ が最初に選ばれる。次に工具径 $D$ とシャンク径 $D_s$ の相関が高い結果であるため、両者のうちでより目的変数に対して単相関の高い変数を選択する。同様に全長 $L$ と刃長 $l$ に対しては単相関の高いどちらかが選ばれる。たとえば、目的変数が軸方向切込み量 $A_d$ のときでは工具径 $D$ と刃長 $l$ が採用される。そこで今度はそのうちのどちらかと表示価格 $PRICE$ でより目的変数に対して相関の高いどちらかを選択し、最終的に工具径 $D$ と相関を比較して高い方を採用する。このように最終的に全ての説明変数間で相関係数0.7以下になるまで削減を行い、3つの説明変数を各条件式に対して選択した。なお変数クラスタ分析と重回帰分析は(株)エスミのExcel多変量解析ver.5.0を用いた。

## 導出された切削条件式の妥当性の検討

➤ 非階層的手法であるK-means法によって工具カタログデータからクラスタ分けし、各クラスタについて階層的手法である変数クラスタ分析を用いて多重共線性のない3つの有意な説明変数を選択し、最後にそれぞれの目的変数に対して選択された説明変数を用いた重回帰分析により、切削条件の決定式を導出した。

# 導出された切削条件式の妥当性の検討

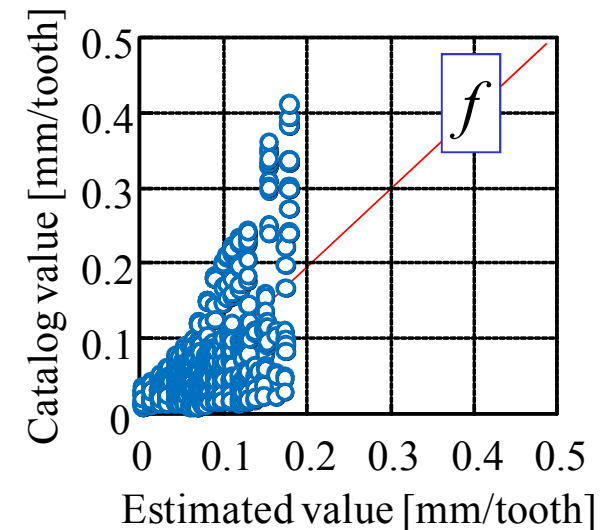
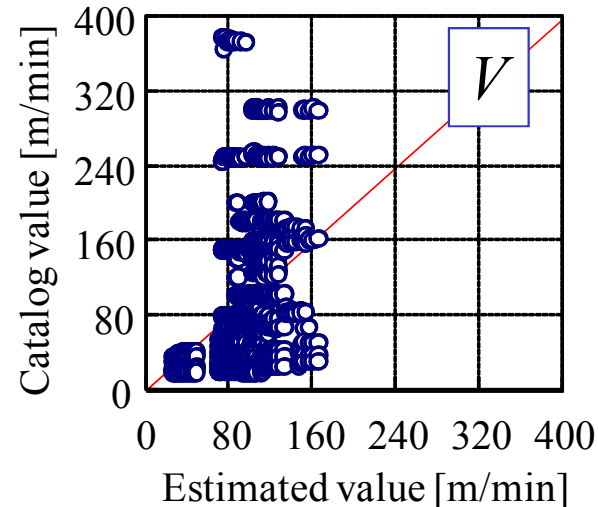
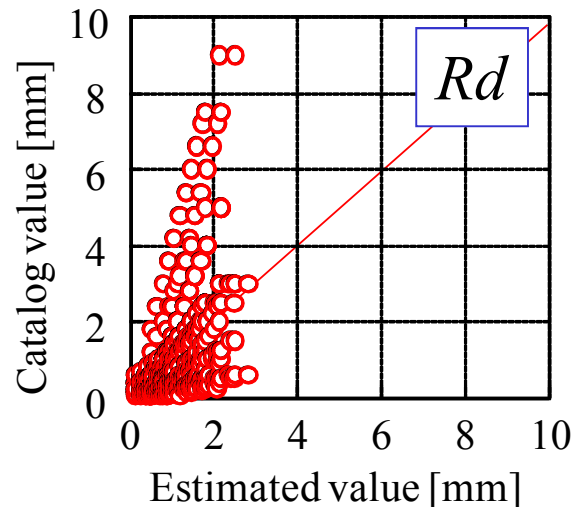
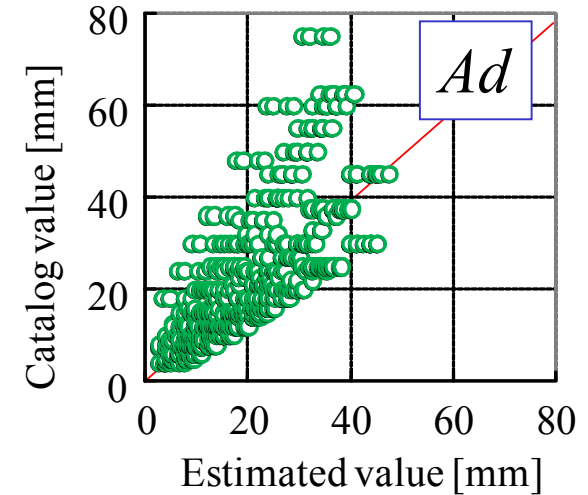
例. Cluster2 側面加工

$$V = 8.76z + 1.49\theta - 1.15HRC + 16.07$$

$$f = 0.0043D + 0.0089z + 0.0034\theta - 0.16$$

$$Ad = 1.62D - 0.12\theta - 0.23HRC + 16.45$$

$$Rd = 0.090D + 0.10z - 0.070HRC - 1.85$$



## 導出された切削条件式の妥当性の検討

Side Processing					Slotting			
Cluster	V	f	Ad	Rd	Cluster	V	f	Ad
1	0.12	0.99	0.87	0.87	1	No data		
2	0.24	0.33	0.67	0.38	2	0.34	0.38	0.75
3	0.41	0.58	0.90	0.68	3	0.27	0.50	0.62
4	0.40	0.87	0.88	0.67	4	0.66	0.64	0.54
5	No data				5	0.85	0.55	0.69
ALL	0.34	0.38	0.49	0.41	ALL	0.13	0.44	0.75

➤ 切削条件決定式の精度を表す決定係数を求めた。ここで表中の ALL は K-means 法によってクラスタリングを行わないでそのまま重回帰分析を行った際の決定係数である。一般的に決定係数が 0.5 以上ならば重回帰分析の精度は良いとみなす。ALL に比べて、非階層・階層型クラスタリング手法を用いた方が、決定係数が 0.5 以上の条件を満たすことが多いことから、併用の効果が表れ、重回帰式としては十分な精度となった。

## 導出された切削条件式の妥当性の検討

- しかし一部側面加工に関して切削速度  $V$  の決定係数が0.5以下のものが残るため、まだ考察の余地が残ることがわかった。
- この結果は、切削速度は切削時の温度、動的荷重、びびり振動などと密接に関係するなど外乱要因が多く、送り速度、切込み量に比べて適切な値の導出が難しいことを示唆している。
- すなわち切削条件の設定の中で、とくに切削速度の選定は最も注意を要することがわかった。

# カタログマイニングプロセス

Selection of variables

Variable cluster analysis  
Multiple regression analysis

Deriving equations for cutting conditions

k-means method

Grouping by tool shape

(a) Selection

(b) Data cleansing

(c) Statistic analysis

(d) Interpretation

Cutting tool catalog data

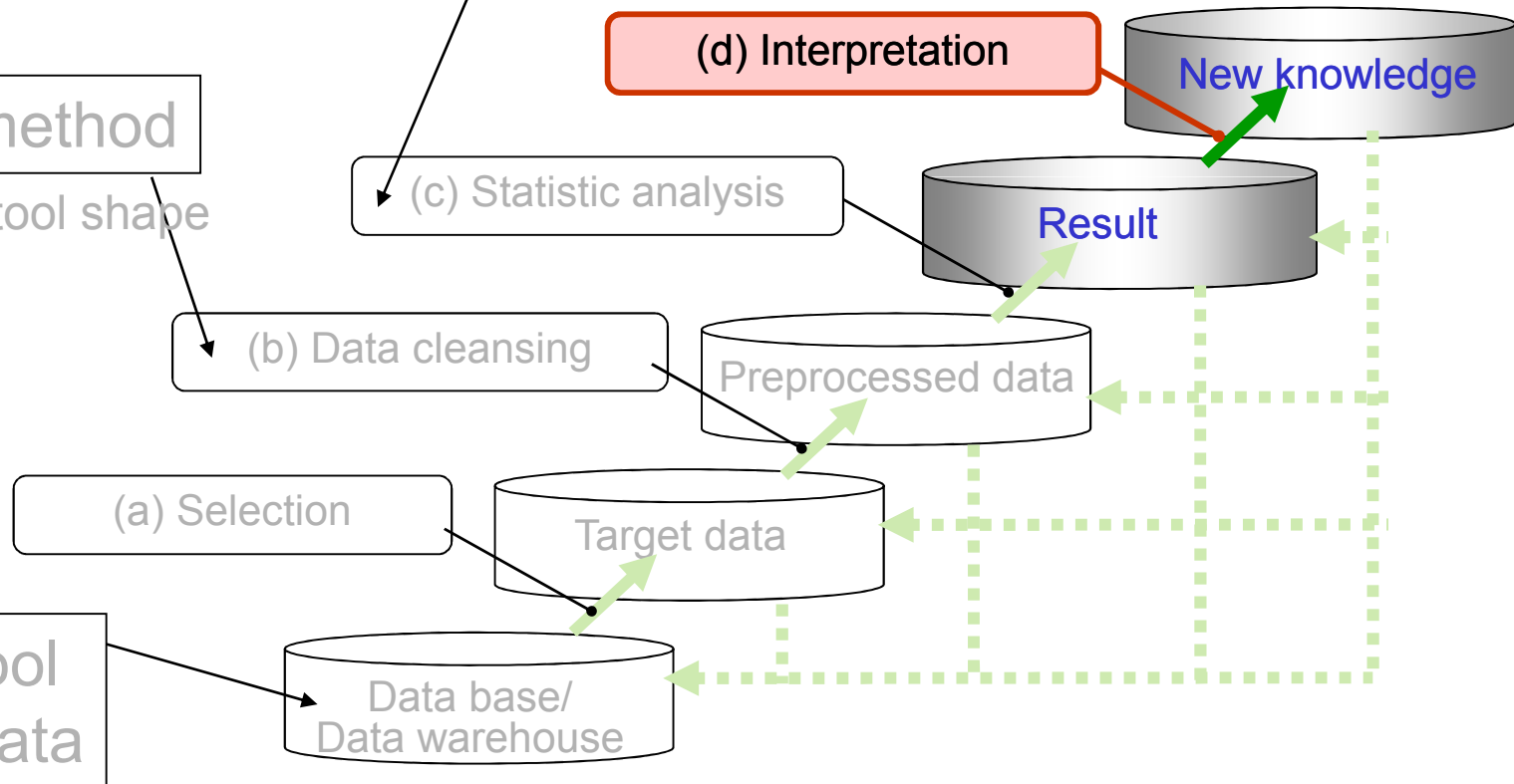
Data base/  
Data warehouse

Target data

Preprocessed data

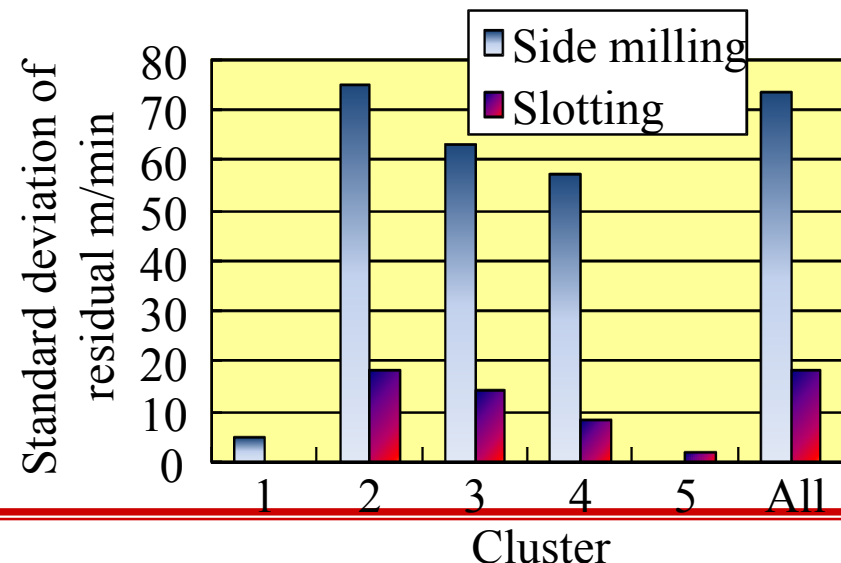
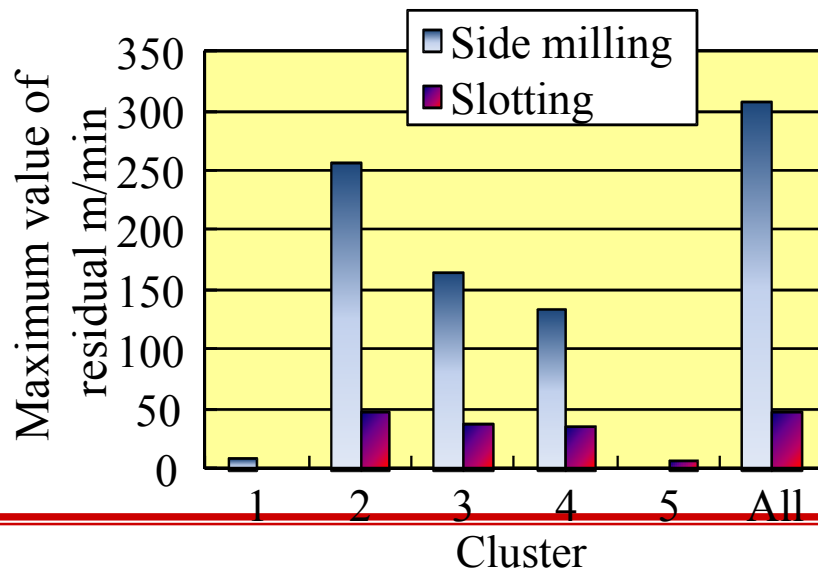
Result

New knowledge



## 知識の解釈

- カタログ値と予測値の誤差を定量的に示すため、導出された条件式の残差の最大値と残差の標準偏差を各目的変数について求めた。目的変数が切削速度 $V$ の場合について、残差の最大値と標準偏差をそれぞれ以下の図に示す
- 非階層・階層クラスタリング手法を併用した効果によって残差は低くなっていることから、本モデルの手法を用いることで、非熟練技能者が初期の目安とするべき程度の条件は十分に示すことができることがわかった。





## 知識の解釈

- 一方で、切削速度 $V$ の条件式において決定係数が0.5未満であるクラスタにおいて、残差の最大値および標準偏差はALLに比べて大きいいため、併用の効果が表れていない
- 精度の良い切削速度 $V$ の条件式を線形式で表すことは困難であるため、非線形な手法による導出の必要性がある点については今後の課題である
- しかし、このような新知識の発掘が行えたのはデータマイニングの効果を示す意味で有効であったと考えられる

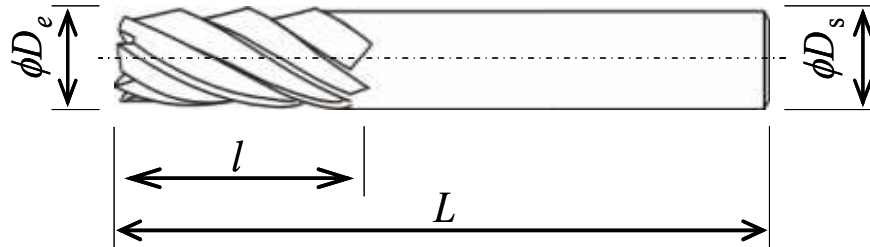
## 結言

- 工具カタログデータに対して、非階層・階層クラスタリング手法を組み合わせたデータマイニング手法を適用し、切削条件を導出できるシステムについて考察した。その結果、非熟練技能者を支援できるシステムとして、工具カタログデータに非階層・階層クラスタリング手法を適用することは、瞬時的な切削条件の導出に有効であることがわかった。

# Appendix

# K-means法の変数選択

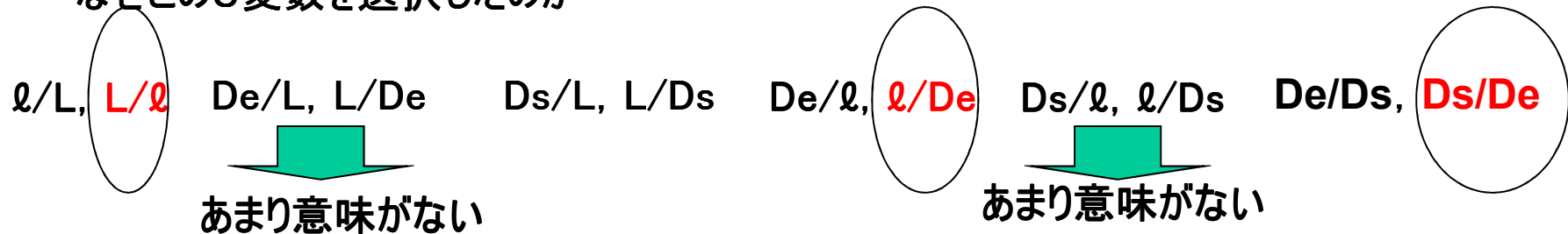
## エンドミル形状モデル



L:全長  $l$ :刃長  $D_s$ :シャンク径  $D_e$ :相当径

・ $L/l$ ,  $l/D_e$ ,  $D_s/D_e$ の3つの係数を導入

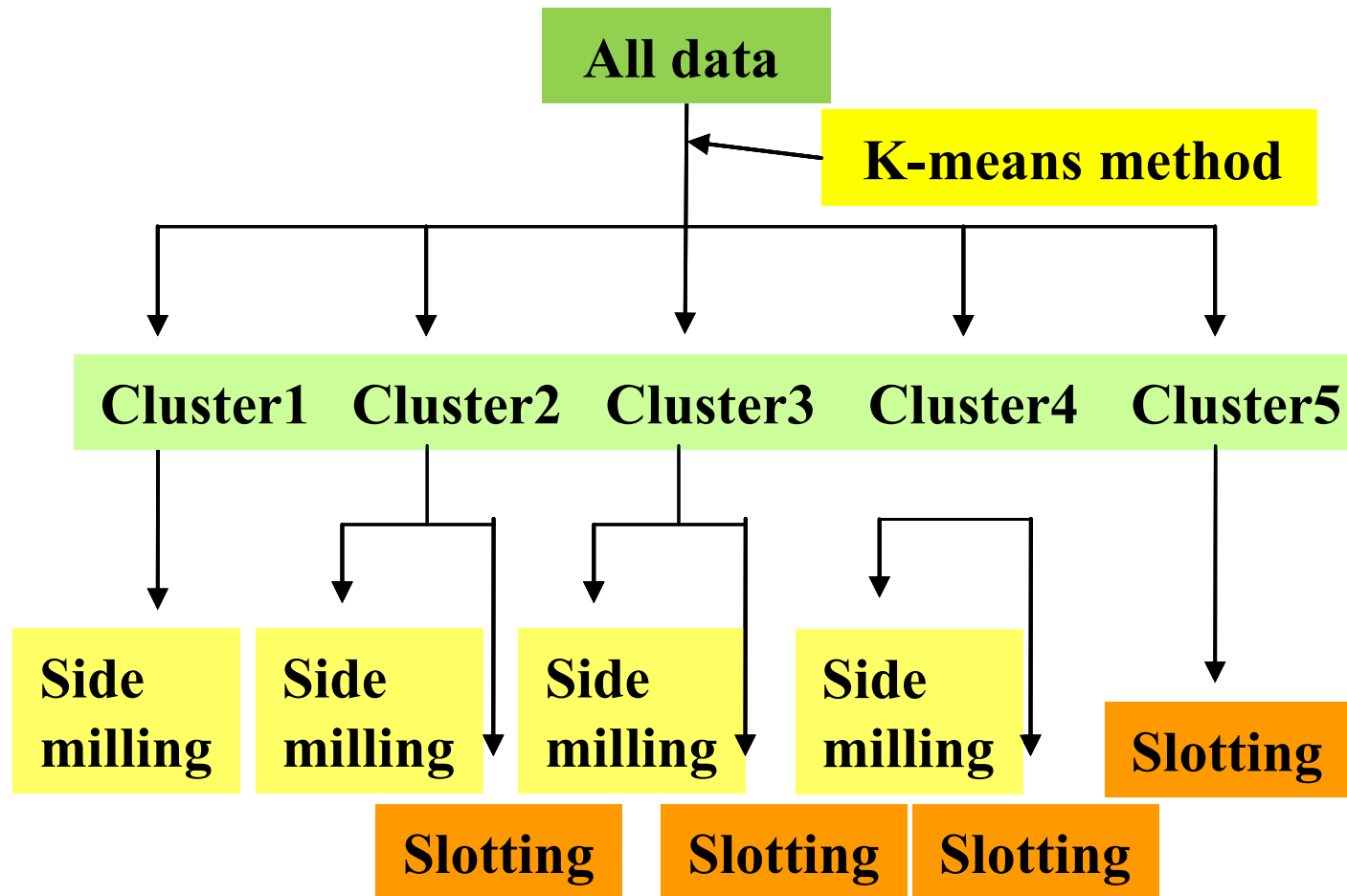
なぜこの3変数を選択したのか



値が1以上になるものを選んだ

比をとることで4変数→3変数とすることができた

# データ分類の状況



各グループそれぞれ溝加工と側面加工に分類

# 相当径 $D_e$ とは

エンドミル→シャンク部に比べて  
刃部は断面積が小さくなる

剛性の低下

〔 相当径 $D_e$ はエンドミル刃部断面積  
を円に置き換えた時の円の直径 〕

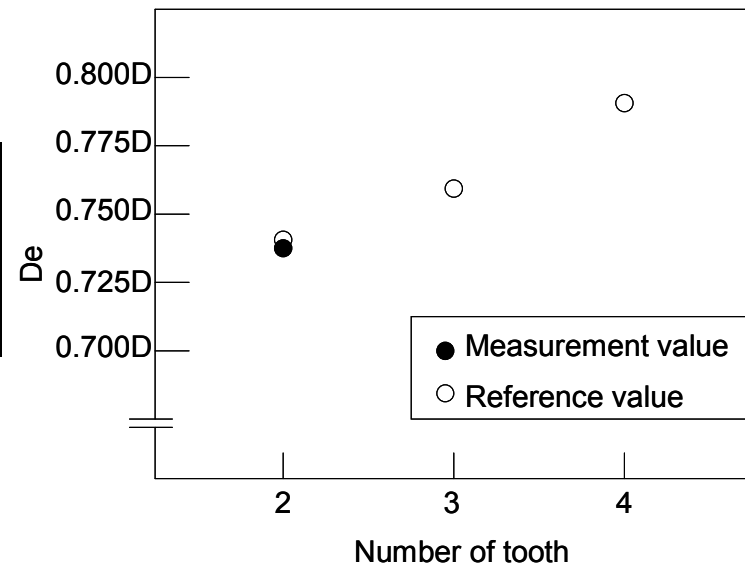
刃数と刃部断面積の関係(文献値)

Number of cutting tooth	Sectional area rate	$D_e$
2	0.55	0.742D
3	0.58	0.762D
4	0.62	0.787D

実際に適用できるか確認するために  
2枚刃, 直径6mmの3本の工具の重さを計測

End-mill	Total length	Edge length	Shanc length	Weight
1	50mm	15mm	29mm	17.4471g
2	50mm	15mm	30mm	17.6323g
3	60mm	22mm	32mm	20.4197g

これより刃部断面積を割り出し,  
相当径 $D_e$ を算出.  
計測した $D_e=0.739$ となった.



# 変数クラスタ分析

なぜ変数クラスタ分析を使うのか？

カタログから抽出した説明変数は、  
切削条件決定に関して全て有効なのか

重回帰分析

相関の高い説明変数が含まれていると正しく解析できない



変数クラスタ分析で相関の高い変数を削減

距離の計算： ウォード法

2つのクラスタ  $C_p$  と  $C_q$  が併合して、  
新たに  $C_k$  というクラスタを生成した場合  
と他のクラスタ  $C_o$  との距離は次式の通り  
となる。

$$d_{ko}^2 = \frac{(n_p + n_o)d_{po}^2 + (n_q + n_o)d_{qo}^2 - n_o d_{pq}^2}{n_k + n_o}$$

(  $n$  はそれぞれのクラスタに含まれる個体数 )