

# 機械学習の一方法とスロッシング波頭衝撃圧 推定式への適用

栗原千鶴子\*・中村秀治\*\*・篠原靖志\*\*\*

基本的な統計的手法と人工知能的なデータ駆動型ヒューリスティクスに基づいて、バラツキの多い数値データから実験式を導き出す機械学習法を検討し、スロッシング波頭衝撃圧の試験データに適用した。その結果、非線形性の強いスロッシング波頭衝撃圧に関して、著者等が先に提案した半理論的評価式と本質的に同等な実験式が機械学習結果からも得られたことを述べている。

**Keywords** : artificial intelligence, machine learning, knowledge acquisition, sloshing impact pressure

## 1. ま え が き

土木構造物の計画、設計、診断等に関するエキスパートシステムが多数構築され、すでに多くの報告がなされてきたが、同時に、工学分野における知識獲得の難しさが再認識され、現状の人工知能技術に対する適正な評価が定着してきたように思われる。

工学知識の多くが、数式や数値解析におけるアルゴリズムという形での表現に適しており、多くの数値解析プログラムに反映されたのに対して、プロダクションルール、フレーム、オブジェクトといった人工知能的知識表現手法に必ずしも適していないことは明らかであり、よりの確な知識獲得、整理、表現法の探求が人工知能技術の当面する最大の課題となっている<sup>1),2)</sup>。すでに、全国レベルでの運用に供されているエキスパートシステム例もあるが<sup>3),4)</sup>、この場合の対象構造物は力学的、かつ、典型的な形状に類別しやすいため、いくつかのパラメータによる有限要素解析モデルの自動生成と、解析結果の有効活用が容易であったことによっている。

知識獲得は外部からの情報入手と、得た情報の体系化より成り、入手した情報を知識にまで高める作業が学習である。学習は非常に多面的な性格をもった人間の高度の精神活動と言える。人工知能研究においても当初からこの精神活動に関心が持たれ、種々の方面から研究されてきた。現在、帰納学習、演繹学習、類推学習の3種類に大別されて、各々の特長を生かした研究が進められている<sup>5),6)</sup>。

帰納学習では、例題が多数与えられた時、それらの間にある一般的な法則性を見出そうとする。例えば、症状

と診断例から診断規則を見つけたりするのがこの例である。一方、演繹学習では知識の論理的判断力は変化しないが、その知識を例題に即してより効率的に活用することを学ぶ。例えば、幾何の証明で公理のみを知っている状態から、応用範囲の広い幾つかの定理を学ぶことがこれにあたる。類推学習では、ある領域での問題に関する知識をもとに、類似ではあるが別の領域の問題に適用できる知識を学ぶ。

学習に関する以上の分類に従えば、実験データから、背後にある物理現象の特性を見出し、あるいは実験式を求める問題は帰納学習の一つとしてとらえることができよう。この種の問題は、統計分野でも古くから研究されてきたが、人工知能分野でのアプローチでは、ヒューリスティクスを重視していることが一つの特長である。研究例としては、LangleyらによるBACONプロジェクトが知られている<sup>7)-9)</sup>。BACONでは、独立変数と従属変数の集合が与えられると、データに基づいてこれらの変数を相互に関係付ける経験則、すなわち、各独立変数をいくつかのオペレータ(+、-、×、÷、2乗など)で組み合わせる少数のデータ駆動型ヒューリスティクス(データに基づく経験的推論規則)を使用して、値がほぼ定数になるような理論項を生成しようとする。具体的には、BACONはまず、一つの独立変数と他の独立変数あるいは従属変数との関係を探し、ある関数関係が見つかると、それを新たな理論項として導入するが、データ値が一定とならない場合は、この理論項も新たな独立変数としてさらに次の関数関係を探し、理論項の生成を繰り返す。最終的に、ほぼ定数となる理論項を生成した時点で終了する。このような方法で、いくつかの物理・化学法則の再発見がなされたという。

しかしながら、既発表文献から判断する限り、BACONの機械学習におけるヒューリスティクス、判断基準は、

\* 正会員 電力中央研究所FBR部主任研究員  
(〒270-11 千葉県我孫子市我孫子1646)

\*\* 正会員 工博 電力中央研究所FBR部座屈研究室長

\*\*\* 正会員 理修 電力中央研究所経済研究所情報システム部



切り分けられたデータ群毎に式の形が全く異なる場合、データ群毎に別個の最終的な式を提示する方向で探索することになるが、相似則が成り立たないような場合を除いて、全てのデータ群で共通の形になるものと考え

(6) 実験式の詳細化

a) 単項式の場合

式(2)のように単項式の場合、

$$y/x_6 = f(x_1, x_2, x_4) \dots\dots\dots (4)$$

とし、

$$\left[ \begin{array}{l} x_1, x_2, x_4 : \text{独立変数} \\ y/x_6 \quad \quad \quad \text{従属変数} \end{array} \right.$$

とみなして、(2)以後の手順を繰り返し適用するか、あるいは、次のデータ駆動型ヒューリスティクスを適用して、実験式を詳細化する。

(ヒューリスティクスー1)

すべての独立変数間の相関が低ければ、従属変数と各独立変数間の相関に注目する。

- ・従属変数  $y$  と独立変数  $x_e$  の相関係数の符号が正ならば、 $y = x_e$  または、 $y = -1/x_e$  となる可能性が高い。
- ・従属変数  $y$  と独立変数  $x_e$  の相関係数の符号が負ならば、 $y = 1/x_e$  または、 $y = -x_e$  となる可能性が高い。

[例] 独立変数  $x_1, x_2, x_4$  間の相関が低く、従属変数  $y$  と  $x_1, x_2, x_4$  との相関係数の符号が、+, +, - であれば、次のいずれかの可能性が高い。

$$y = \frac{x_1 \cdot x_2}{x_4}, \quad \frac{x_2}{x_1 \cdot x_4}, \quad \frac{x_1}{x_2 \cdot x_4},$$

$$-x_1 \cdot x_2 \cdot x_4, \quad \frac{1}{x_1 \cdot x_2 \cdot x_4}, \quad \frac{x_1 \cdot x_4}{x_2}, \quad \frac{x_2 \cdot x_4}{x_1}, \quad \frac{x_4}{x_1 \cdot x_2}$$

(ヒューリスティクスー2)

特定の独立変数間に正の相関が極めて高く、他の相関が低い場合、独立変数同志の比を新たな独立変数として、相関の高い独立変数を除外する。

また、特定の独立変数間に負の相関が極めて高く、他の相関が低い場合、独立変数同志の積を新たな独立変数として、相関の高い独立変数を除外する。

この操作の後、ヒューリスティクスー1を適用する。

(ヒューリスティクスー3)

データが切り分けられていない場合、ヒューリスティクスー1, -2の適用により、全データに対して定数に近い値をとる  $f$  の形が見出されたならば、それを実験式とする。

(ヒューリスティクスー4)

データが切り分けられている場合、各データ群の中で定数に近い値をとる  $f$  の形が共通して見出されたならば、それを実験式の一部とする。さらに、各データ群で生じた定数と、切り分けて一時的に除外していたデータを用いて、同様の手順を繰り返し、最終的な実験式の形

を見出す。

b) 複数項より構成される場合

式(3)のように複数項より構成される場合、

$$y = ax_6 + bx_3 \dots\dots\dots (5)$$

または、

$$y = ax_6 + bx_3 + c \dots\dots\dots (6)$$

として重回帰分析し、偏回帰係数により  $a, b, c$  を定めれば、

$$f(x_1, x_2, x_4) = a \text{ (定数)} \dots\dots\dots (7)$$

$$g(x_3, x_5) = b \text{ (定数)} \dots\dots\dots (8)$$

となる。回帰式が定数項  $c$  を有するか否かは、問題ごとに判断すべき性質のものである。

式(1)のように切り分けられている場合、最大6組の組み合わせに対して、式(5)または式(6)の  $a, b$  が次のように定められる。

$$\left[ \begin{array}{l} a = a_{i_1 j_1}, a_{i_1 j_2}, \dots, a_{i_6 j_2} \dots\dots\dots (9) \\ b = b_{i_1 j_1}, b_{i_1 j_2}, \dots, b_{i_6 j_2} \end{array} \right.$$

(7) 実験式の係数の決定

以上の手順で、単項あるいは複数項からなる実験式の形が決定したならば、回帰分析により各項の係数を決定する

(8) 探索の終了判定

以上で良い実験式が得られなかった場合、他のデータと比べて極端に実験式との差が大きい少数のデータは削除して、重相関係数、偏回帰係数を求めなおして再評価する。なお、実験式として不十分であれば、(2)以下のステップにもどり、前回と異なる設定で探索を繰り返す。その場合、理論項の自動発生等において、現象に関する他の予備知識が有用になる。探索を繰り返すことにより、十分良い実験式が見出された時点で終了する。

以上の探索手順をフロー図で示すと、図-1の通りである。

3. 波頭衝撃圧の実験条件と実験結果

前述のBACONが、その有用性の確認例として示した式は、ごく単純な理想気体の法則、オームの法則等にすぎない。本機械学習法の適用例としては、先に提案した平屋根に対する地震時波頭衝撃圧の評価式を取り上げ、その有用性について考察する。

実験条件の概要を以下に述べる。

表-1に示す小型、中型、大型模型を用い、長周期大振幅加振装置によりスロッシング実験を行なった<sup>14)~16)</sup>。加振条件は、表-2のスロッシング1次固有振動数における定常正弦波加振(最大30波)とした。

波頭衝撃圧は、液面の乱れが少ない衝突第1波目に最大圧が現れると言われており、本実験でも、上部空間がある程度以上あれば、波頭衝撃圧の最大値は、衝突第1波目から1.5(~2)波目に現れた。



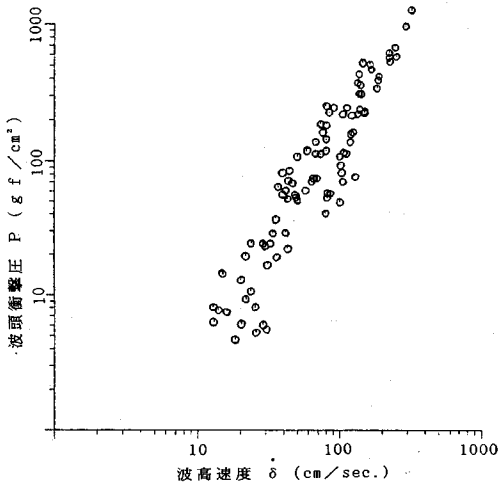


図-3 波高速度と波頭衝撃圧の関係

$d$ : 加振振幅 [cm]  
 $\delta$ : 波高速度 [cm/sec.]

である(図-2)。

データ数は、表-3に示す通り、93個とする(前論文<sup>10</sup>)においては、上部空間の非常に小さいデータも含めていたが、スロッシング波高の十分成長する以前の状態であるためここでは除外した。また、図-3には、小型、中型、大型模型実験におけるスロッシング衝突波高速度と衝撃圧の関係を図示した。

#### 4. 本機械学習法により得られた波頭衝撃圧の実験式

2. で述べた機械学習法に従い、表-3のデータから波頭衝撃圧の実験式を定める手順および結果を示す。

##### (1) データの切り分け

表-3から明らかな通り、 $R$ (容器半径)、 $h$ (水深)、 $H$ (上部空間高さ)、 $\omega$ (円振動数)については、共通の値が多く現れ、7つのデータ群に切り分けることができる。

##### (2) 主たる独立変数の次数の推定

データ群毎に、

$\log_{10} P$  : 従属変数  
 $\log_{10} d, \log_{10} \delta$  : 独立変数

として、ステップワイズ回帰分析を行い、主たる独立変数の次数を推定した結果は、次の通りである。

データ群番号	次数
1	$\delta^{1.93}$
2	$\delta^{1.86}$
3	$\delta^{1.72}$
4	$\delta^{1.76}$
5	$\delta^{1.59}$
6	$\delta^{1.74}$
7	$\delta^{1.21}, d^{0.44}$

したがって、7番のデータ群を除いて、 $\delta$ が主要な独立変数であり、各群共通して2乗まで考慮すればよいと判断できる。

##### (3) 新たな理論項の自動発生

下記の5つの理論項を自動発生させる。

$d, \delta$	$\delta^2$
$d \delta$	$d \delta^2$

##### (4) ステップワイズ回帰分析の適用

新たな理論項に対してステップワイズ回帰分析を行い、重要な理論項を推定した結果は、次の通りである。

データ群番号	重要な理論項(左から)
1	$\delta^2, \delta$
2	$\delta^2, d \delta, d, \delta$
3	$\delta^2, d \delta, \delta, d$
4	$\delta^2, \delta$
5	$\delta^2, \delta$
6	$\delta^2, \delta$
7	$\delta^2, d$

##### (5) 実験式の形の大まかな決定

(4)の結果から、 $d$ (加振振幅)が重要な理論項とされる場合のあることを認めつつも、各データ群に共通する実験式の形としては、次式を選定できる。

$$P = (c_0 + c_1 \delta) \delta^2 \dots \dots \dots (10)$$

##### (6) 実験式の詳細化

式(10)を式(11)のように変形し、各データ群毎に回帰分析した結果を、 $R, h, H, \omega$ と共に示すと下表の通りである。

$$P/\delta^2 = c_0 + c_1 \delta \dots \dots \dots (11)$$

データ群番号	$c_0$	$c_1$	$R$	$h$	$H$	$\omega$
1	0.1506	0.0123	27.5	14.	6.	6.10
2	0.4264	0.0148	27.5	16.	4.	6.35
3	0.2818	0.0075	111.5	63.	37.	3.55
4	0.8196	0.0085	111.5	73.	27.	3.67
5	1.4591	0.0090	111.5	83.	17.	3.76
6	1.0147	0.0099	327.0	192.	52.	1.85
7	1.1552	0.0170	327.0	211.	33.	1.90

また、 $R, h, H, \omega$ の相関行列は以下の通り。

	$R$	$h$	$H$	$\omega$
$R$	1.0000			
$h$	0.9942	1.0000		
$H$	0.8202	0.7814	1.0000	
$\omega$	-0.9210	-0.9266	-0.9017	1.0000

独立変数の間の相関がいずれも高いので、2. に述べたヒューリスティクスを単純には適用できないが、すでにデータ数が少なく数値自体にある程度特性があるので、この段階では統計的手法よりAI的手法の方が有利である。

すなわち、相関の高い独立変数同士の積または商を求め、 $c_0, c_1$ との相関を考察すれば、 $c_0$ に対しては、 $(c_0/(R/h)), (c_0/(h/R)), (c_0/(R/H)), (c_0/(H/R)), (c_0/(R*\omega)), (c_0/(h*\omega)), (c_0/(H*\omega))$ のうちで、最も定数に近づくものを探せばよいことになる。同様に、 $c_1$ に対しても、 $(c_1/(R/h)), (c_1/(h/R)), (c_1/(R/H)), (c_1/(H/R)), (c_1/(R*\omega)), (c_1/(h*\omega)), (c_1/(H*\omega))$ のうちで定数に近づくものを探せばよいことになる。

結果は次の通りである。

群番号	$c_0 \cdot \frac{h}{R}$	$c_0 \cdot \frac{R}{h}$	$c_0 \cdot \frac{H}{R}$	$c_0 \cdot \frac{R}{H}$	$\frac{c_0}{R*\omega}$	$\frac{c_0}{h*\omega}$	$\frac{c_0}{H*\omega}$
1	.077	.296	.033	.691	.0009	.0018	.0041
2	.248	.733	.062	2.941	.0024	.0042	.0168
3	.159	.499	.094	.849	.0007	.0013	.0021
4	.537	1.251	.199	3.387	.0020	.0031	.0083
5	1.086	1.961	.223	9.599	.0034	.0046	.0228
6	.596	1.729	.161	6.382	.0017	.0029	.0105
7	.745	1.781	.117	11.44	.0019	.0029	.0184

群番号	$c_1 \cdot \frac{h}{R}$	$c_1 \cdot \frac{R}{h}$	$c_1 \cdot \frac{H}{R}$	$c_1 \cdot \frac{R}{H}$	$\frac{c_1}{R*\omega}$	$\frac{c_1}{h*\omega}$	$\frac{c_1}{H*\omega}$
1	.0063	.0242	.0027	.0564	.000073	.000144	.000336
2	.0086	.0254	.0022	1.021	.000085	.000146	.000583
3	.0042	.0133	.0025	.0226	.000019	.000034	.000057
4	.0056	.0130	.0021	.0351	.000021	.000032	.000036
5	.0067	.0121	.0014	.0592	.000021	.000029	.000141
6	.0058	.0169	.0016	.0623	.000016	.000028	.000103
7	.0110	.0264	.0017	1.683	.000027	.000042	.000271

$c_0$ に対しては、いずれの場合も定数に近づけられたとは言えないが、 $c_0/(h*\omega)$ が最良で、 $c_0/(R*\omega)$ も次によい。 $c_1$ に対しては、 $c_1/(R/H)$ がよい。

(7) 実験式の係数の決定、探索の終了判定

回帰分析により、実験式の係数および重相関係数を求めると、

$$P = 0.0022 \frac{R}{H} \delta^2 + 0.0022 h \omega \delta \dots \dots \dots (12)$$

(重相関係数：0.9948)

あるいは、

$$P = 0.0023 \frac{R}{H} \delta^2 + 0.0013 R \omega \delta \dots \dots \dots (13)$$

(重相関係数：0.9943)

となる。

両式とも、表-3のすべてのデータに対する重相関係数が十分高く、探索の過程であえて2者択一したような箇所も無かったので、本データに対する探索は終了する。

(8) 考察

(2)~(7)においては、統計的手法に依存する割合が高かったが、工学的問題に対する現実的対応の一面を示したものとえよう。バラツキのある多数のデータから知識を学習するに際して、AI的手法と基本的な統計的手法を適切に組み合わせるのは、極めて効果的である。

波頭衝撃圧の現象に関しては、山本<sup>17)</sup>の先駆的な研究があり、衝撃圧は波高速度の2乗に比例した式を提案している。その後の他の研究も実験・実測データの大きなバラツキと、波高速度の2乗と1乗の間にくる傾向を認めつつも、2乗則を支持してきた。一方、浅井<sup>18)</sup>は実験データのみから波高速度の1.6乗則を提案した。これらの実験的事実は、スロッシング波頭衝撃圧に関して、波高速度の1乗に比例した影響因子の存在することを暗示している。文献10)においては、種々のスケールによる実験結果より、若干安全側に評価して次の半理論式を提案した。

$$P = \frac{6.63 \rho}{g} \cdot \frac{(H+h)R}{H} \cdot \frac{(0.2R+H+h)}{(0.4R+H+h)^2} \delta^2 + \frac{0.35 \rho}{g} \cdot \frac{R(H+h)}{(0.4R+H+h)} \omega \delta \dots \dots \dots (14)$$

(ここで、 $g$ ：重力加速度)

式(13)は、波頭衝撃圧は波高速度の2乗項と1乗項の和で表されるとする上記半理論式(14)と本質的に同じものである。また、表-3のデータに限定すれば、 $R$ と $h$ の相関は極めて高いので、式(12)とも同じといえる。従って、今回の例から、単なる数値データの機械学習により、何らの物理的考察なくとも、的確な実験式にたどり着く可能性のあることは明らかである。

5. 結 論

以上、基本的な統計的手法と、AI的な少数のデータ駆動型ヒューリスティクスに基づいて、数値データから実験式を導き出す機械学習法を検討した。検討内容の多くはアブダクション (Abduction：発想的推論) であり、明確な論拠に欠けるものの、工学問題に対して実際に実行可能な方法を提案したものである。適用例としては、非線形性の強いスロッシング波頭衝撃圧現象を取り上げ、その適用性を確認した。

得られた成果を要約すると次の通りである。

(1) コンピュータの利用技術を高度化し、人工知能的な利用を進めるためには種々の学習法を確立する必要がある。数値データからの的確に実験式を推定するのその一つであるが、常に多くの不確定要素を含んだ工学的

問題に対して、機械学習法確立の可能性が明らかになった。

(2) 提案する機械学習法は、少数のデータ駆動型ヒューリスティクスと、相関分析、ステップワイズ回帰分析などの統計的手法を組み合わせたものである。パラツキのある多数のデータから有用な情報を抽出する過程で統計的手法が有効であり、ある程度条件が絞りこまれた後、式の形を詳細検討する際にAI的なヒューリスティクスが有効である。

(3) 非線形性の強いスロッシング波頭衝撃圧現象に対して、半理論的に導かれた評価式と本質的に同等な実験式が機械学習からも得られた。本式の特長は、波高速度の2乗項のみでなく、1乗項をも含むものである。

なお、現段階でのコンピュータ・プログラムは、既存の統計解析プログラムとの関係上、FORTRANで作成されている。今後、一つの機械学習システムまで高めていくには、適用例を参照しつつ種々の閾値を適切に設定し、理論項の自動発生法をさらに高度化し、式の形の選定に係わる判断ルールを整備して、一部AI言語で書き直す必要も生じるであろう。それらは今後の課題である。

#### 参考文献

- 1) 上野晴樹・石塚 満：知識の表現と利用，知識工学講座2，オーム社，1989.
- 2) 中村秀治・寺野隆雄：土木構造物エキスパートシステム，オーム社，1987.
- 3) 中村秀治・松浦真一・寺野隆雄・篠原靖志：水力鋼構造物の寿命予測エキスパートシステムとその適用，土木学会論文集，No.374/I-6，1986.
- 4) 松本幸雄・中村秀治・山本広祐：水力鋼構造物評価診断システムの試験的利用について，電力土木，No.226，1990.7.
- 5) 大須賀節雄・佐伯 胖：知識の獲得と学習，知識工学講座3，オーム社，1988.
- 6) 篠原靖志・矢沢利弘：知識獲得のための機械学習，電力中央研究所報告，Y87016，1988.
- 7) Cohen, P.R. and Feigenbaum, E.A. Eds.: The Handbook of Artificial Intelligence III, William Kaufmann Inc., 1982.
- 8) Langley, P., Bradshaw, G.L. and Simon, H.A.: Rediscovering Chemistry with BACON System, in Michalski, R.S. et. al. Eds.: Machine Learning: An Artificial Intelligence Approach, Tioga, 1980. (邦訳 BACONによる化学法則の再発見，学習と問題解決，知識獲得と学習シリーズ第2巻，共立出版，1987.)
- 9) Langley, P., Bradshaw, G.L. and Simon, H.A.: The Search for Regularity - Four Aspects of Scientific Discovery, in Michalski, R.S. et. al. Eds.: Machine Learning: An Artificial Intelligence Approach, Morgan Kaufmann, 1983. (邦訳 規則性の探求，発見的学習，知識獲得と学習シリーズ第6巻，共立出版，1988.)
- 10) 栗原千鶴子・増子芳夫・桜井彰雄：固定屋根を有する液体タンクのスロッシング波頭衝撃圧に関する検討，土木学会論文集，No.432/I-16，1991.
- 11) 奥野忠一・久米 均・芳賀敏郎・吉沢 正：多変量解析法，日科技連，1978.
- 12) 奥野忠一・芳賀敏郎・矢島敬二・奥野千恵子・橋本茂司・古河陽子：統 多変量解析法，日科技連，1978.
- 13) 芳賀敏郎・橋本茂司：回帰分析と主成分分析，統計解析プログラム講座2，日科技連，1980.
- 14) Yashiro, T., Kawashima, H., Nakagawa, M., Sakurai, A., Masuko, Y. and Kurihara, C.: A study on non-linear sloshing of pool type LMFBR coolant, Trans. of the 9th Int. Conf. on Structural Mechanics in Reactor Technology, Vol.E, pp.459-464, 1982.
- 15) Sakurai, A., Masuko, Y., Kurihara, C., Ishihama, K., Yashiro, T., Chang, Y.W. and Rodwell, E.: Seismic sloshing experiments of large pooltype fast breeder reactors, Nuclear Engineering and Design, 113, pp.423-433, 1989.
- 16) 増子芳夫・栗原千鶴子・萩原 豊・沢田義博：スロッシング波頭衝撃圧の実験的研究，電力中央研究所報告，研究報告：U88059，1989.
- 17) 山本善之：地震による石油タンクの液面の動揺と衝撃圧力，高圧力第3巻第1号，1965.
- 18) 浅井 修，他：固定屋根を有する円筒液体タンクの耐震設計（第2報），圧力技術，第17巻第4号，1979. (1990.12.18受付)

## A METHOD OF MACHINE LEARNING AND ITS APPLICATION TO THE ESTIMATION OF SLOSHING IMPACT PRESSURE

Chizuko KURIHARA, Hideharu NAKAMURA and Yasusi Sinohara

To build knowledge-based systems more effectively, it is required to improve the performance of the acquisition of knowledge, especially in engineering domain, the acquisition of quantitative knowledge. One of the most promising candidates is the application of machine learning techniques to (semi-) automatic acquisition of quantitative knowledge.

The purpose of this paper is to propose an inductive machine learning method for the formulation of experimental equations from noisy empirical data. It combines statistical techniques and heuristics to reduce the search space of possible equations effectively. As an example of its application, the estimation of sloshing impact pressure at flat roof surface of tanks is taken up, which one of the authors reported previously.