

相関と因果



因果をとらえ、モデル化する 必要性は何か？

- 時々刻々かわる環境
- ダイナミクスを記述するモデルがない予測システムでは信頼性評価が困難
 - ✓ 定量的リスク評価
 - ✓ 突発的現象の確率的予報

巨大なシステムの制御(第八代所長 赤池先生らの業績)

多変量自己回帰モデル y_n は観測ベクトルデータ

$$y_n = \sum_{j=1}^m A_j y_{n-j} + w_n$$



【写真提供】太平洋セメント(株)

セメント生産プロセスへの統計的方法の適用は、秩父セメント株式会社で初めて実現され、複雑なシステムに対しても制御の可能性を明らかにしました。

(Nakamura and Akaike, 1981)



【写真提供】九州電力(株)

統計的方法を用いた火力発電所のボイラー温度制御は、九州電力株式会社で世界に先駆けて実用化され、優れた制御性を示しています。

(Nakamura-Akaike, 1981)

船の制御

船舶のオートパイロット

(第十代所長 北川先生らの業績)



【写真提供】東京商船大学(現 東京海洋大)

統計的な方法によって設計されたオート・パイロットの研究
成果が生かされ、更に新しい研究が進められています。
(Ohtsu-Horigome-Kitagawa, 1978)

$$\mathbf{x}_n = F\mathbf{x}_{n-1} + G\mathbf{v}_n, \quad \mathbf{y}_n = H\mathbf{x}_n$$

状態空間モデル

$$J(\mathbf{x}_n) = \sum_{j=1}^L \left(\mathbf{x}_{n+j}^T Q \mathbf{x}_{n+j} + r_{n+j-1}^T R r_{n+j-1} \right),$$

評価基準

$$\mathbf{r}_n = \mathbf{y}_n - H\mathbf{x}_{n|n-1}$$

オートパイロットシステムは、業界トップシェアを誇っています。

国内シェア: 50%

海外シェア: 38%

日本郵船、商船三井、川崎汽船をはじめとする船会社が
当社のお客様です。

出典 <http://rois.ac.jp/outline/pdf/saisyukougi.pdf>
<https://job.rikunabi.com/2016/company/top/r101900017/>

新幹線強風予報システムの開発に協力

(第十代所長 北川先生 & 第11代所長 樋口の業績)



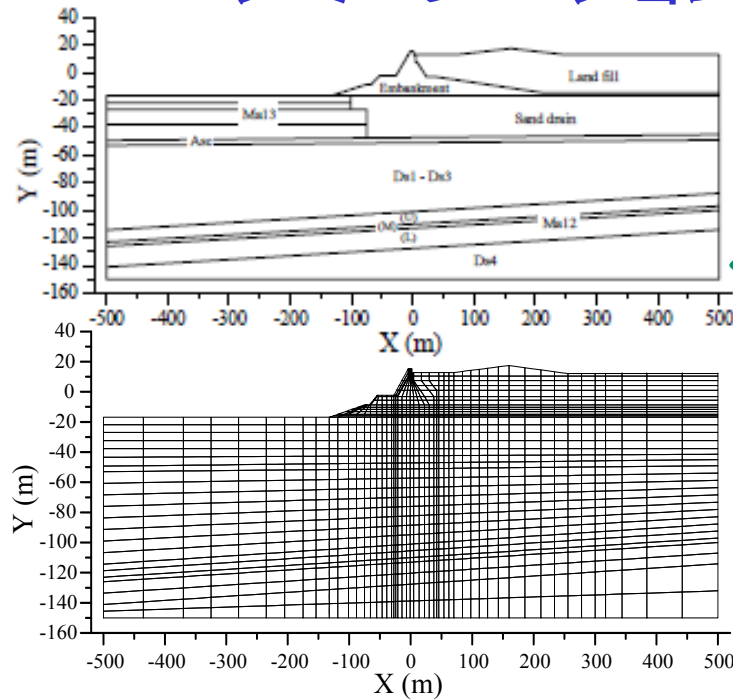
状態空間モデル の利用



朝日新聞平成14年2月17日(日) 14版第2社会面

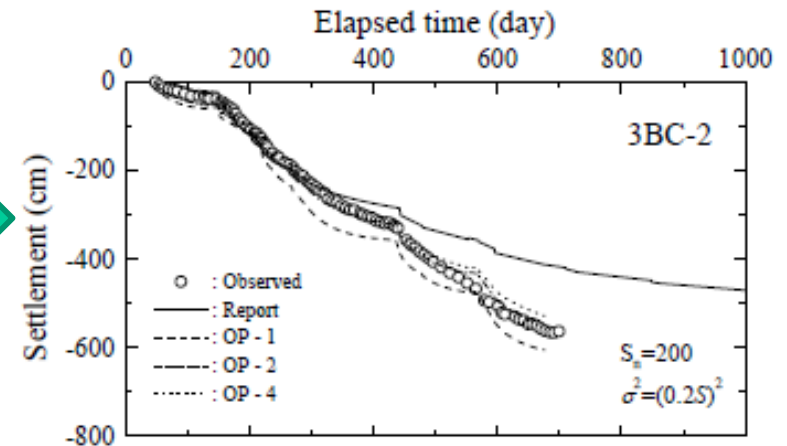
データ同化例：神戸空港・地盤沈下

シミュレーションモデル(支配方程式)との統合



地盤変形モデル

データ同化



沈下量データ

- 直接見ることができない地中の土の状態がわかる
- 予測精度の向上で、中途での工法変更が可能に

(村上・藤澤(京大), 珠玖・西村(岡大)各氏との共同研究)

ビッグデータになるとさらに 何が大変か？

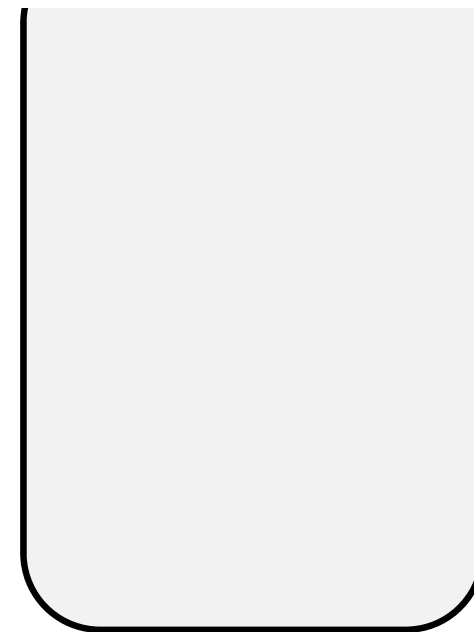
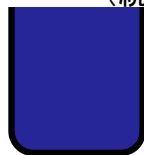
- ビッグデータの解釈・利用には前処理は必要不可欠。さらにその処理に多くの労力と時間が費やされる
- ネットワーク型データ特有の空間相関の統計処理のスキルが必須

価値（←目的による！）密度の定義

$$\text{価値密度(目的)} = \frac{\text{価値総量(目的)}}{\text{データ総量 Volume}}$$

ビッグデータは
「価値密度」がかなり低い

（統計数理研究所 丸山教授（現PFI）； PFI 岡野原氏 談）



抽象物を対象にすると 誰もプロセスに分解できない！

Data cleansing, Data Editing, Data Curating

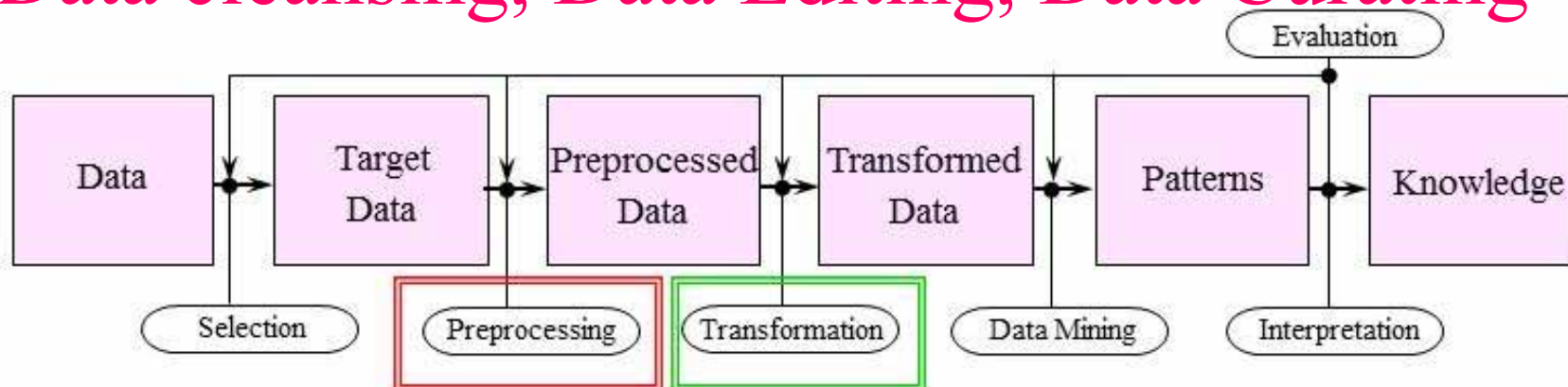


図 1: Fayyad等による知識発見のプロセス

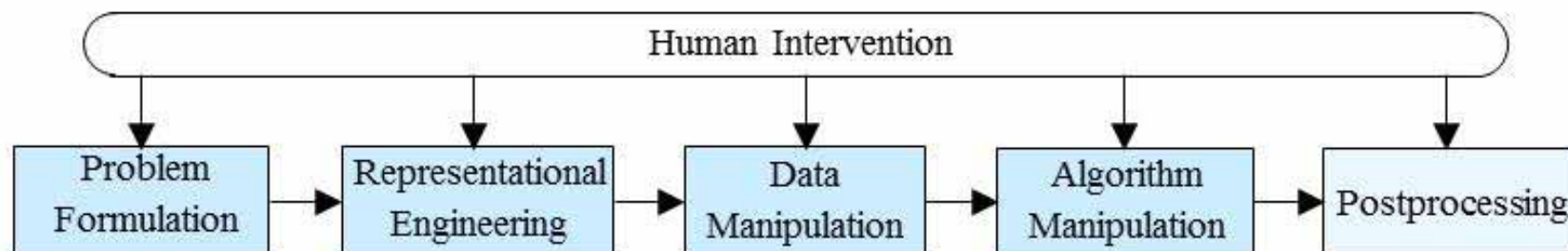


図 2: Langleyによる知識発見のプロセス

データ処理の流れ

発見科学研究 成果報告集
宮野悟(平成13年3月)