可視化の世界

計算科学が拓く世界

2015年1月7日（水）

国際高等教育院
小山田耕二

内容

- 可視化について
- 可視化事例
- 視覚的分析環境
- 課題

本日のテーマと目的

- 可視化は、計算機や計測装置等から生成される膨大な数値データから気付きを得るための基盤技術として重要になっている。本講義では、計算科学と密接な関係にある可視化技術の基礎と応用について説明する。

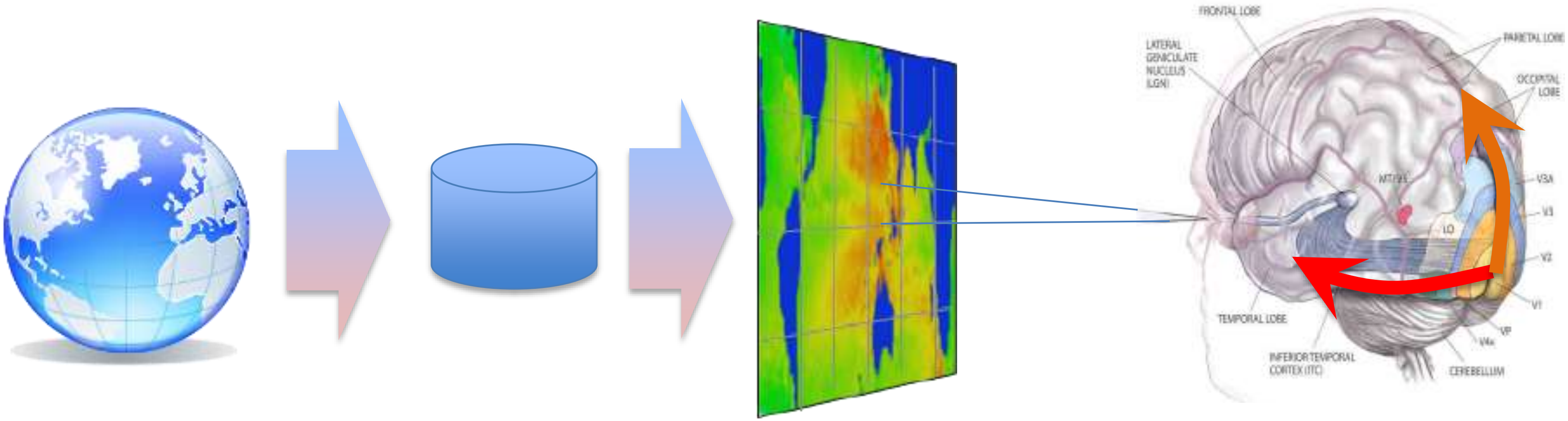
気付きを促す可視化

可視化について

可視化って何かな？

見えないものを見える
ようにすること

See the unseen



現象　　データ　　画像

ビッグデータ時代

Source: IDC's Digital Universe Study, sponsored by EMC, June 2011

年間1.8ゼタバイトのデータ生成

1.8ゼタバイトってどれくらい？

- ゼタ(zetta):

1,000,000,000,000,000,000,000=$1{,}000^7$=10^{21}

- 575億のiPAD(32GB)に格納されたデータ

世界最速　京コンピュータ

http://www.nsc.riken.jp/K/diary.html

TOP500リストに登録した「京」のシステムは、現在整備途中段階のもので、672->864筐体（CPU数68,544個->88,128個）の構成です。LINPACK（リンパック）ベンチマークでは、世界最高性能の8.162->10.51ペタフロップスを達成し、TOP500リストの首位を獲得しました。また、実行効率は93.0%と高水準の記録を達成しました。 日本のスーパーコンピュータがTOP500リストで第1位となるのは、2004年6月以来のこととなります。

順位	システム名称	設置場所	ベンダー	国名	LINPACK演算回数（テラFLOPS）
1	K computer	理研　計算科学研究機構	Fujitsu	日	10,510
2	天河1A号	天津スパコンセンタ	NUDT	中	2,566
3	Jaguar	オークリッジ研	Cray	米	1,759
4	Nebulae（星雲）	深圳スパコンセンタ	Dawning	中	1,271
5	TSUBAME2.0	東京工業大学	NEC/HP	日	1,192
6	Cielo	ロスアラモス研（サンディア研）	Cray	米	1,110
7	Pleiades	NASA・エイムズ研究センタ	SGI	米	1,088
8	Hopper	ローレンス・バークレイ研	Cray	米	1,054
9	Tera-100	原子力庁（エネルギー研）	Bull	仏	1,050
10	Roadrunner	ロスアラモス研	IBM	米	1,042

データと情報

データ

何かを文字や符号、数値などのまとまりとして表現したもの

情報

データを人間が解釈したもの

データを認識させること

データ

何かを文字や符号、数値などのまとまりとして表現したもの

可視化

情報

データを人間が解釈したもの

気付きを促す可視化

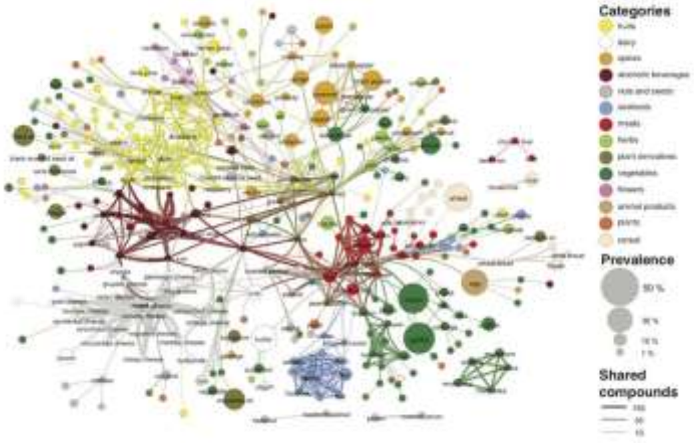
いろいろな可視化

計測データの可視化

流れの可視化
（データ提供：Gustavo R.S. Assi）

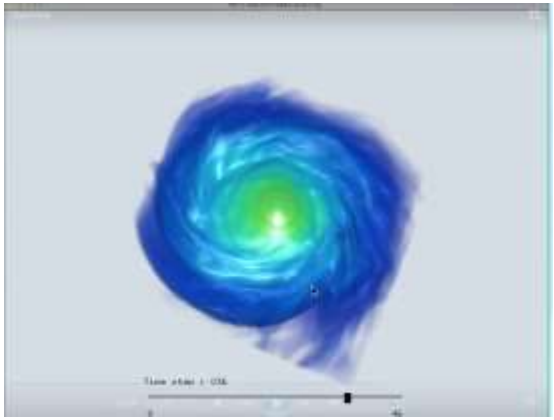
船鉾の可視化
（データ提供：田中先生＠立命館大学）

医療データの可視化
（データ提供：www.oagpubco.com）



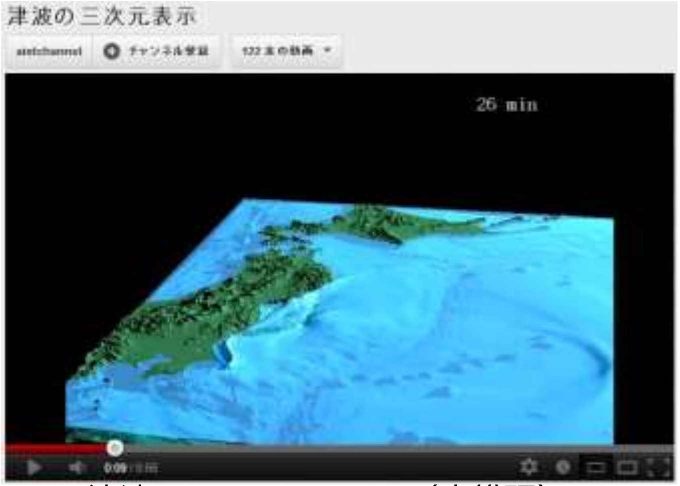
Yong-Yeol Ahn, et. al,
Flavor network and the principles of food pairing, Nature

シミュレーション結果の可視化

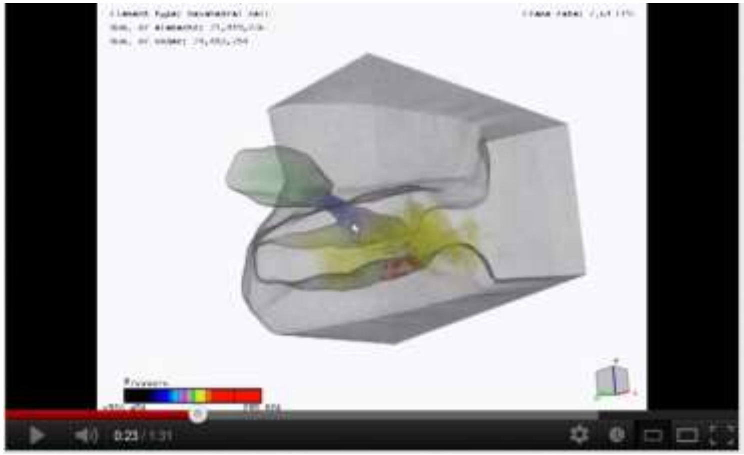


ブラックホールシミュレーション
（木内先生@京都大学）

自動車衝突解析(LOCTITE sitio Web www.loctite.com)

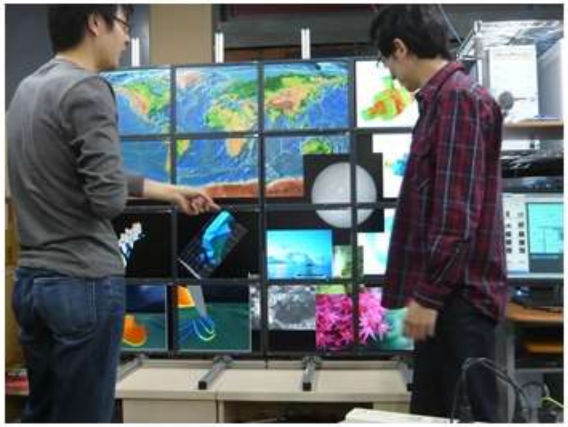


津波シミュレーション（産総研）



大規模口腔流体解析結果（野崎先生@大阪大学）

大型表示装置を使った可視化



複数アプリケーションの表示例

星間物質の乱流シミュレーション結果
（村主先生@白眉プロジェクト）

立体映像提示（アナグリフ方式）
（キッズサイエンススクール）

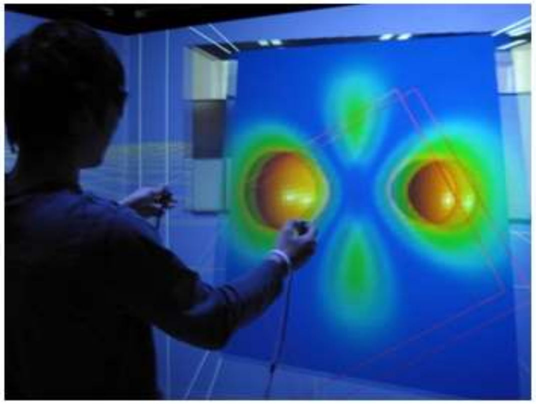
N体シミュレーション結果（暗黒物質）
（矢作先生@京大メディアセンター）

大規模ポンプの構造解析結果
（奥田先生@東京大学）

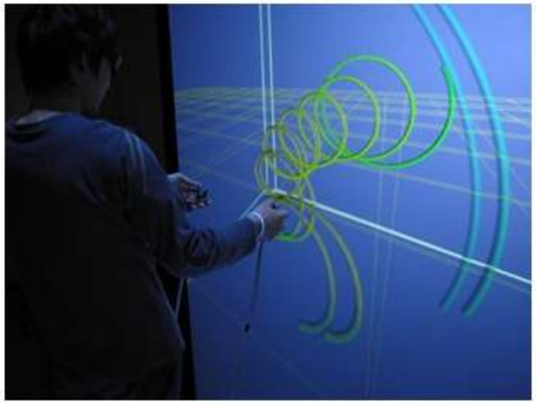
システム構成

PC x 11
NIC: 1Gbps
LCD: 40

没入表示装置を使った可視化

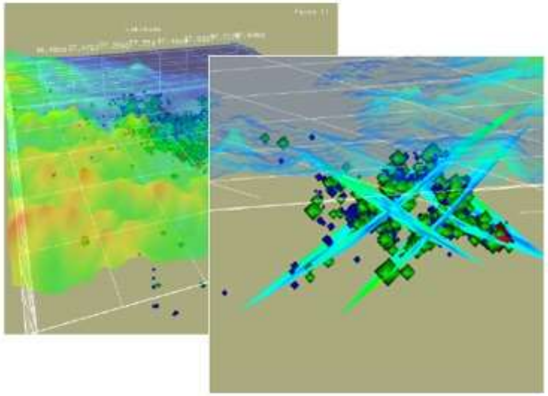


任意断面可視化



流線可視化

没入仮想空間提示



新潟中越地震の震源分布（断層面の推定）
（片尾先生@京大防災研）

- PC x 4
- 磁気装置
- 偏光メガネ
- 操作デバイス

システム構成

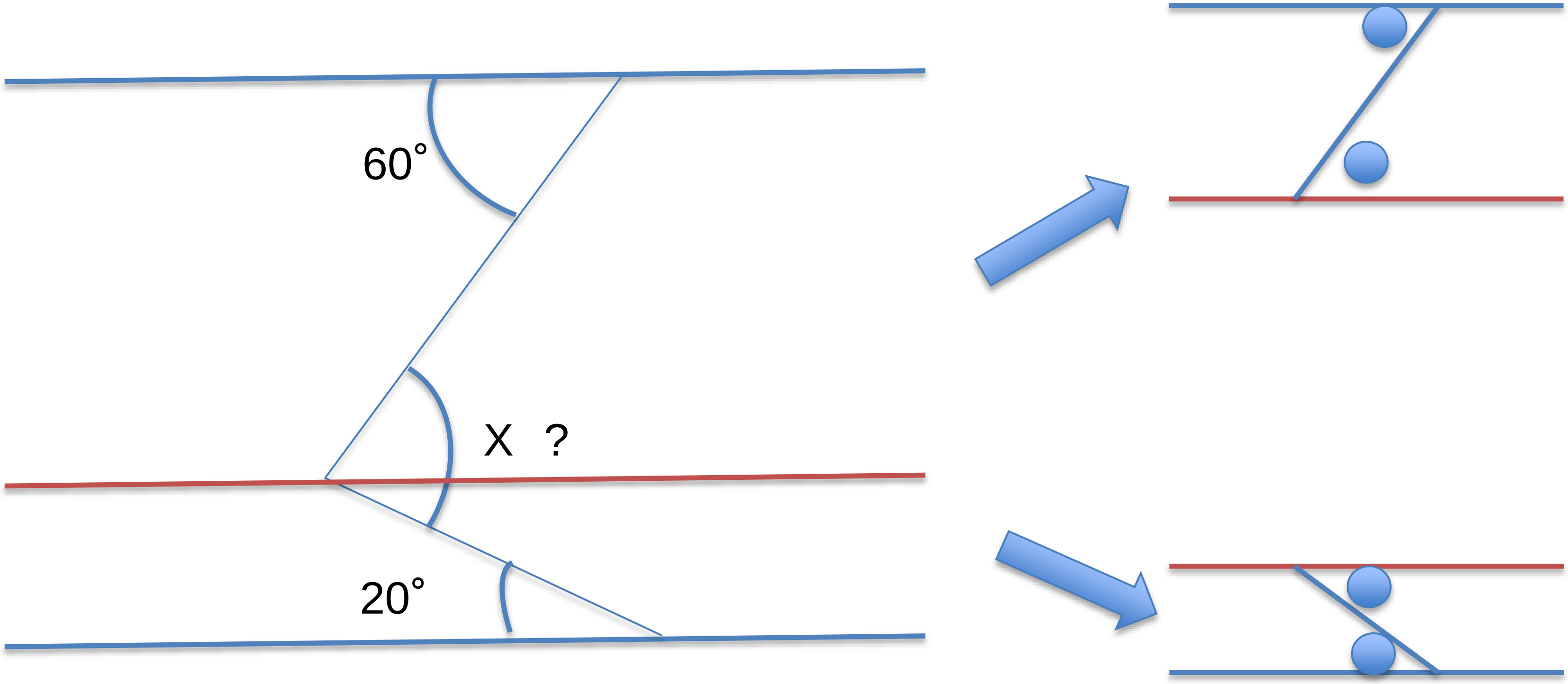
可視化の世界

視覚的分析環境について

The science of analytical reasoning facilitated by interactive visual interfaces
(NVAC Illuminating the Path, 2005)

幾何の問題を考えよう

- 角度を計算せよ

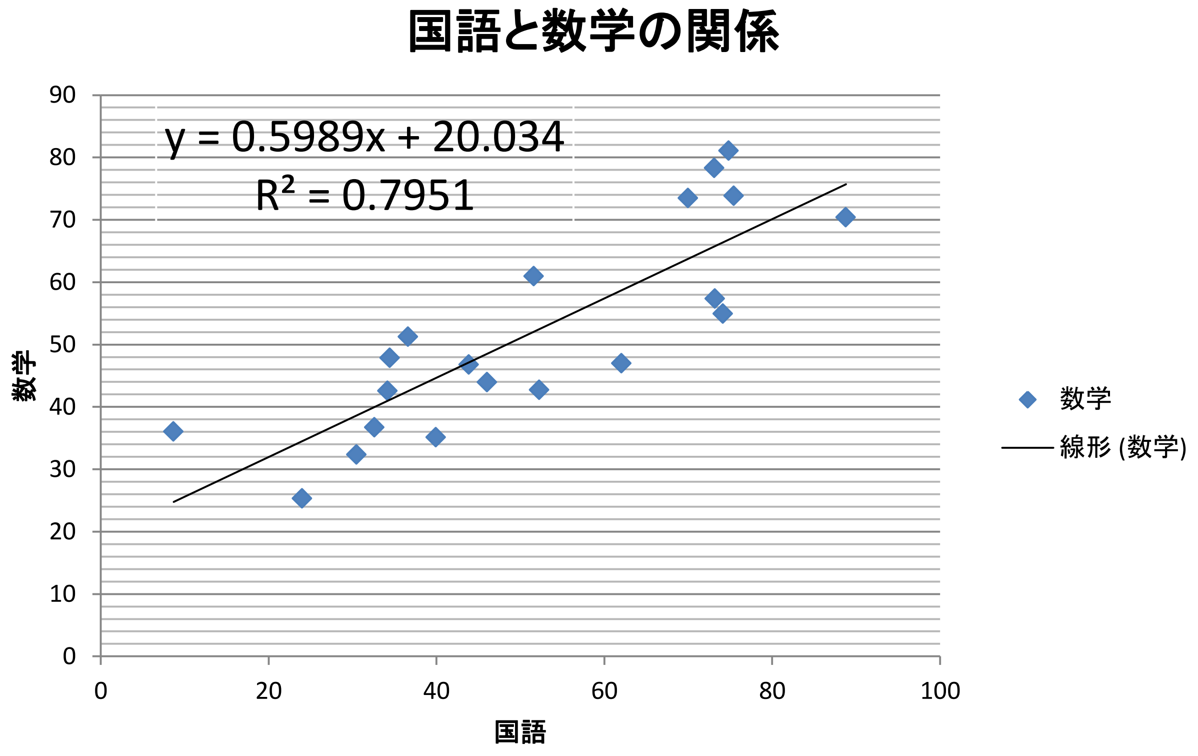


既知のパターンを呼び起こす

国語と数学の関係を調べよう

- 国語と数学の点数の分布を可視化しなさい

国語	数学
45.5432	51.03428
59.91986	61.83581
80.29279	58.51433
74.29664	74.61355
52.10229	47.2127
22.77132	40.57847
19.54681	35.75629
39.00053	40.58302
71.56791	73.91526
87.17153	88.06188
51.62499	57.34637
45.00857	56.99122
77.30697	64.01505
83.82299	82.66693
61.65687	61.47212
89.75105	70.3592
47.71503	46.61471
94.53155	87.83143
45.04513	51.98799
0.421381	16.93079



因果関係の科学的発見

政府コメントにおける「科学」

1. 2012年3月11日細野大臣は、フジテレビ系列の報道番組で、「科学的に安全が証明されているがれきの受け入れ拒否は、被災地の切り捨て」と言い切った。
2. 2012年5月19日細野大臣は、ストレステストの結果について、記者団に 「科学的知見に基づき作った判断基準は十分説明できた。 各首長の反応については真摯に受け止めたい」 と語った。
3. 2012年8月31日橋下徹大阪市長は、大阪府と大阪市が開いた震災がれきの受け入れの説明会で反対意見が相次いだことについて「反対の人しか集まっていない。事実誤認で、現在の一般的な科学論から離れた見解が多い」と市役所で記者団に述べ、批判した。

科学的方法

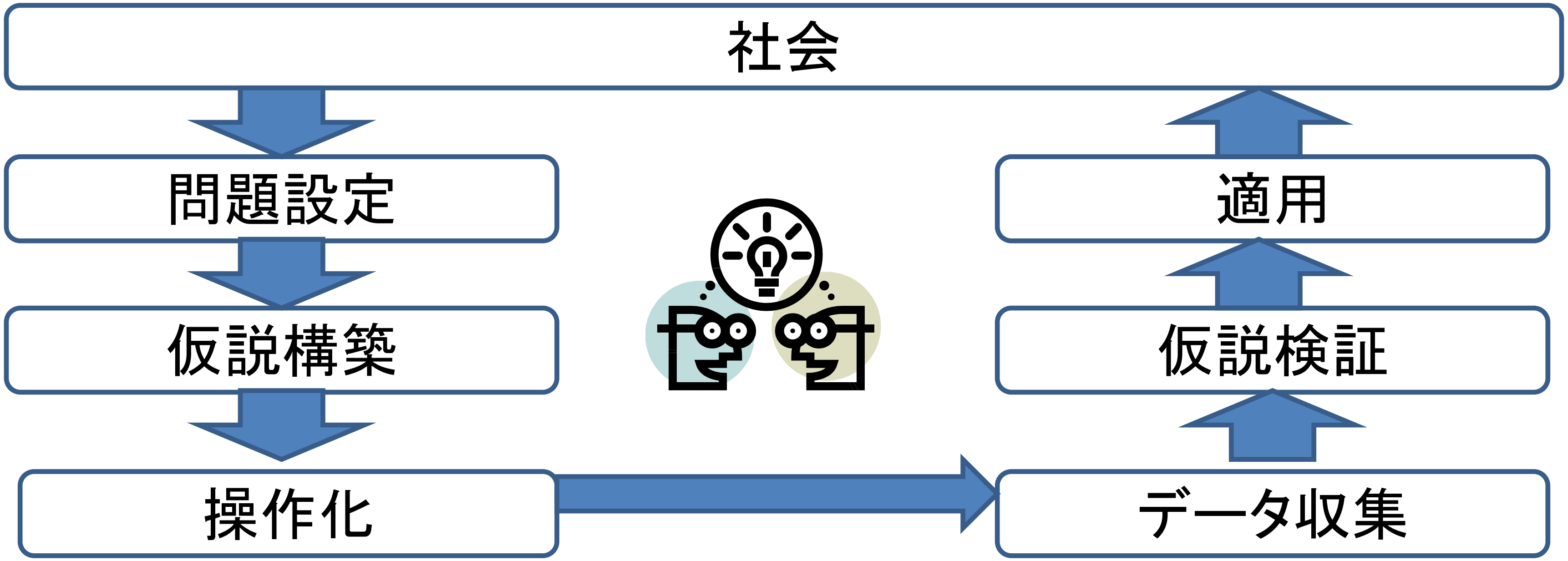
F. James Rutherford and Andrew Ahlgren, Science for All Americans , 1989

1. 物事を調査し、結果を整理し、新たな知見を導き出し、知見の正しさを立証するまでの手続きであって、（仮説検証）
2. その手続きがある一定の基準を満たしているもののことである。（査読）

仮説検証法

坂下昭宣：社会科学方法論

- もともと自然科学における方法論であったが、社会科学においても採用されるようになった
- 因果法則の定式化とその実証
  - 現象の因果関係を仮説により説明
  - 仮説の真偽を経験的データにより実証



仮説検証法の例

- アイデア
  - 大学において，仲間に恵まれて，楽しいクラスに所属すると授業成績が高い
- 仮説化（変数概念への変換）
  - 授業が面白く、一緒に学習を進める仲間に恵まれていると感じるクラスでは，受講学生の成績平均値が高い
- 概念操作化（調査項目への変換）
  - 授業への興味（原因：説明変数）
  - 真面目さ（原因：説明変数）
  - 仲間との絆（原因：説明変数）
  - 熱意（原因：説明変数）
  - 授業成績（結果：被説明変数）

授業成績＝4.6*興味+3.0*絆-10.2

視覚的分析環境について

The science of analytical reasoning facilitated by interactive visual interfaces
(NVAC Illuminating the Path, 2005)

計算科学シミュレーションと仮説検証

- アイデア
  - 自動車の車体は、力がかかって熱くなっているところで破壊が進みやすい
- 仮説化（変数概念への変換）
  - 車体上で応力と温度ともに高くなっているところでは，車体のクラックが多く発生する
- 概念操作化（調査項目への変換）
  - 応力（原因：説明変数）
  - 歪み（原因：説明変数）
  - 温度（原因：説明変数）
  - クラック発生率（結果：被説明変数）

クラック発生率＝3.2*応力+1.0*温度+1.2

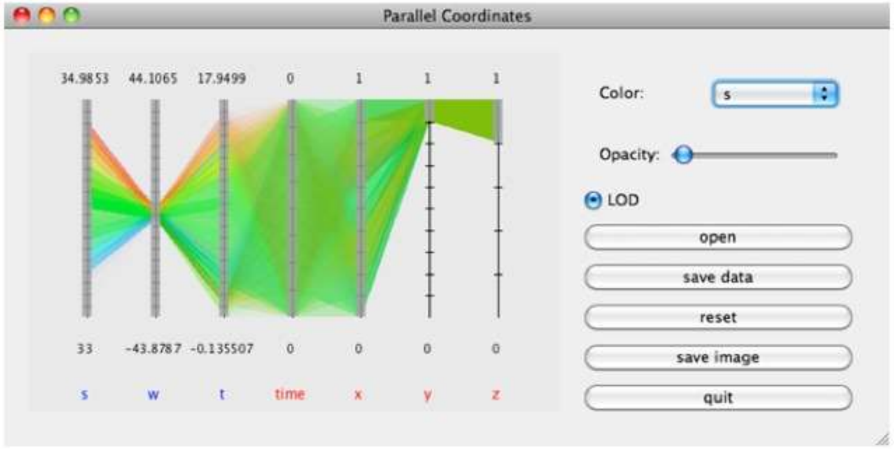
計算科学シミュレーションと仮説検証

- アイデア
  - さんまの漁獲量は、小規模渦の周辺で、塩分を多く含む暖かい領域で多くなる
- 仮説化（変数概念への変換）
  - 釣り機あたりの漁獲高は、微細渦の中心で、塩分濃度と水温がともに極大となる場所で、最大化する
- 概念操作化（調査項目への変換）
  - 微細渦中心からの距離（原因：説明変数）
  - 塩分濃度勾配（原因：説明変数）
  - 水温勾配（原因：説明変数）
  - 釣り機あたりの漁獲高（結果：被説明変数）

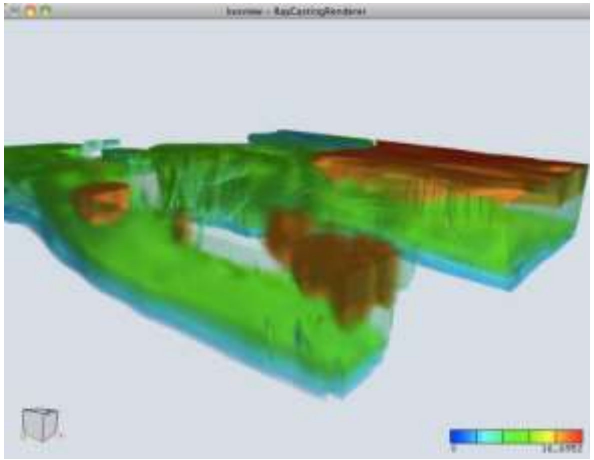
漁獲高＝-4.6*塩分濃度勾配-3.1*水温勾配+72.3

視覚的分析環境への要求

- 仮説構築・検証を支援するために…
  - 統合可視化：座標空間と変数空間との連動
  - 融合可視化：多変数間の関係

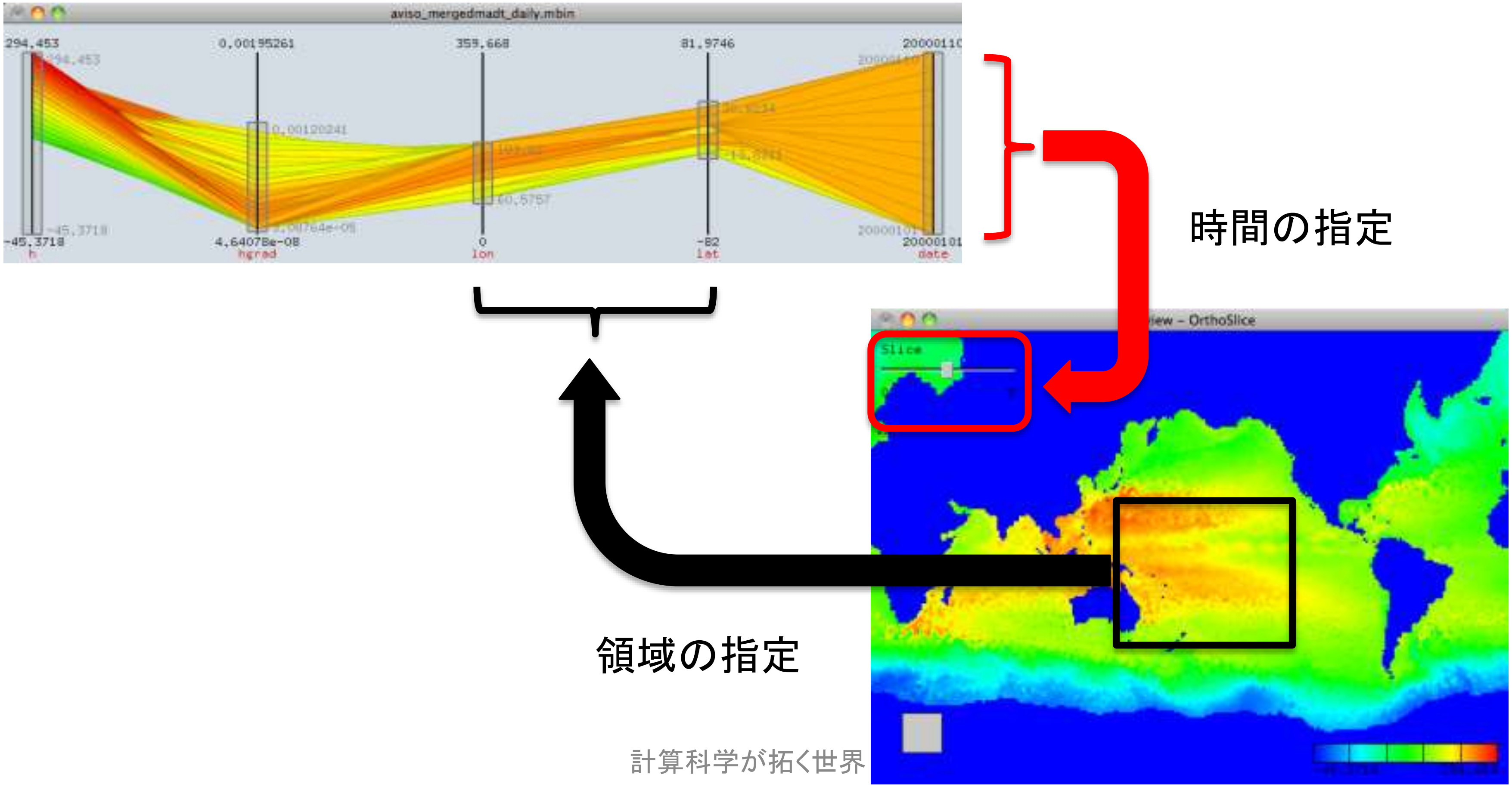


＋



解析領域の指定

- 緯度軸、経度軸の範囲指定



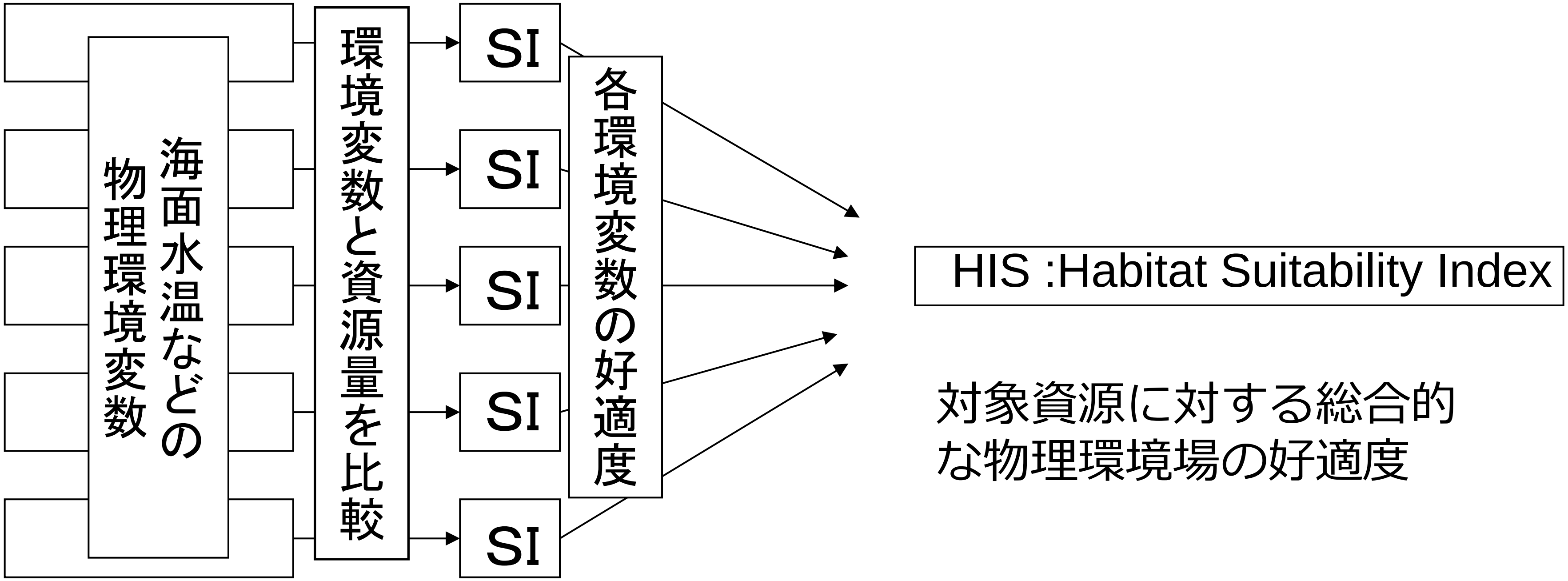
可視化の世界

ピンポイント漁場予測

HTTP://WWW.MEXT-ISACC.JP/_PUBLIC/EVENT2011_DATA/POSTER/RECCA_AWAJI.PDF

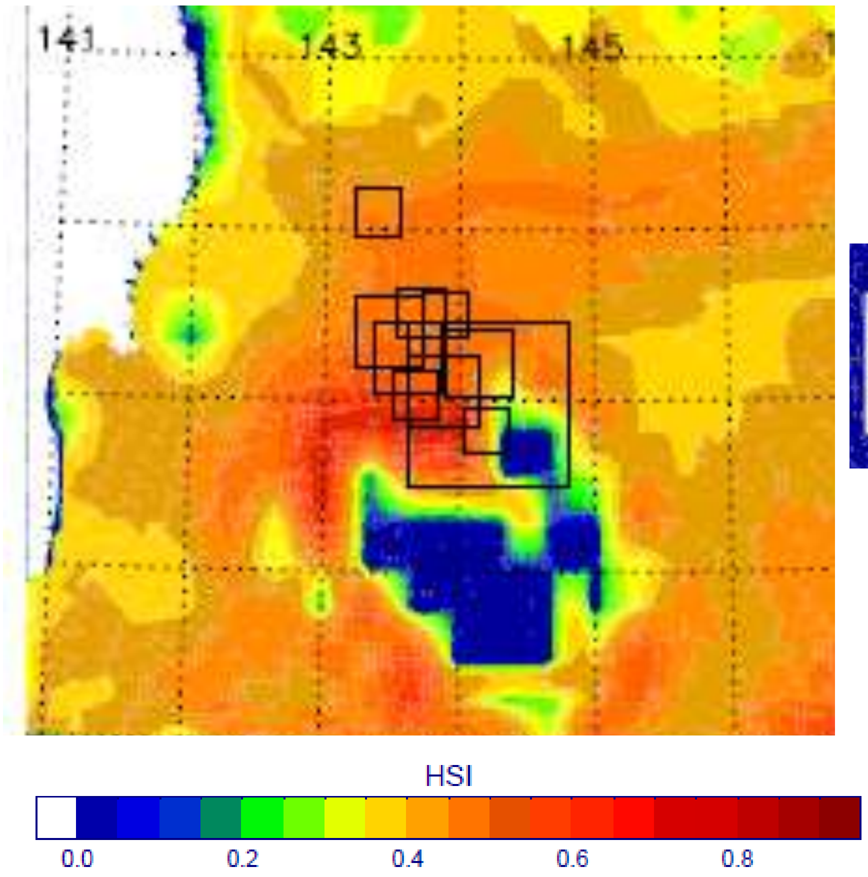
HSI統計予測モデル

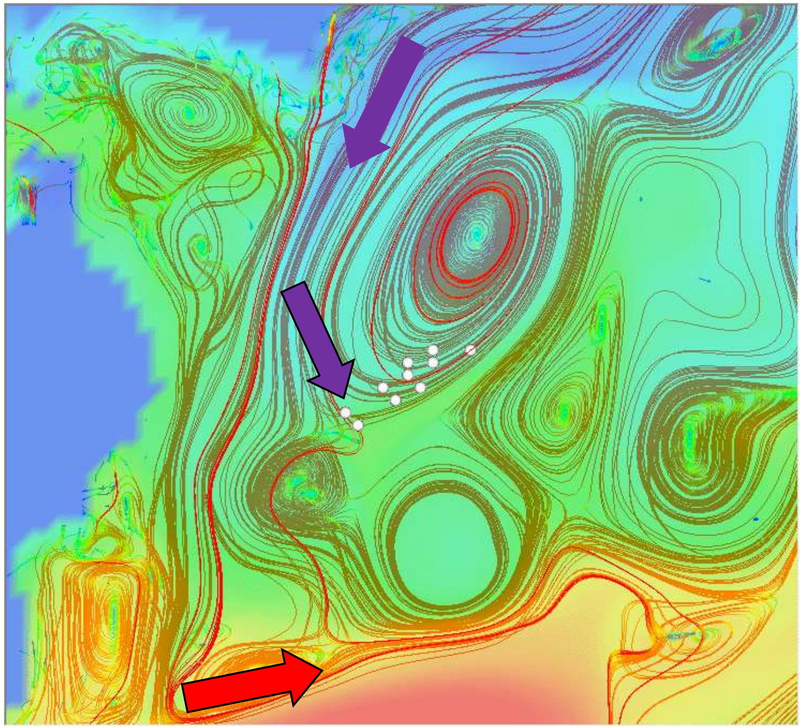
・評価対象の資源量と海洋環境変数の関係を統計的に示す手法。
・対象資源にとってどの様な海洋環境場が好適であるかを推定することが出来る。



高精度漁場探索

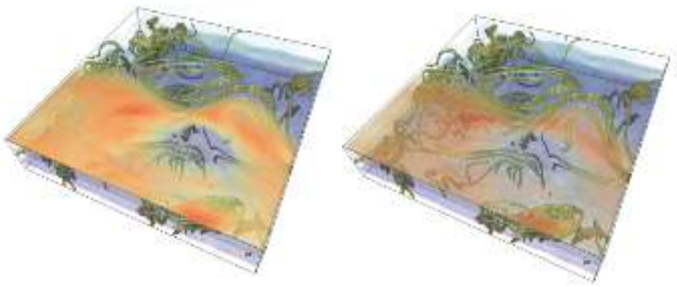
http://www.mext-isacc.jp/_public/event2011_data/poster/RECCA_awaji.pdf

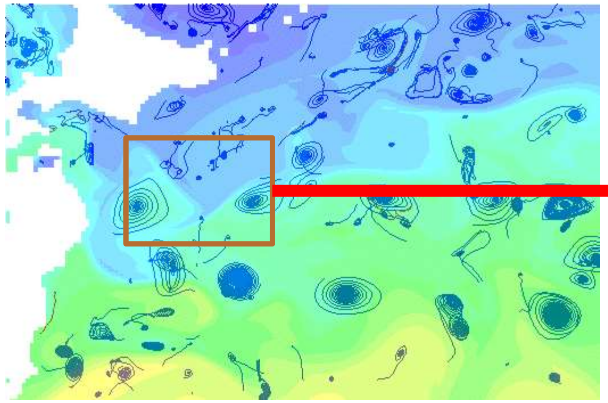




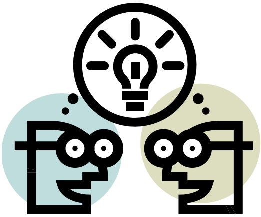
- CPUE data is obtained from fishermen at Feb. 2007.
- HIS(Habitat Suitability Index) is calculated as F(T,V)

大気・海洋結合データ同化プロダクトについて、
気づき・ひらめきを引き出す可視化環境の整備

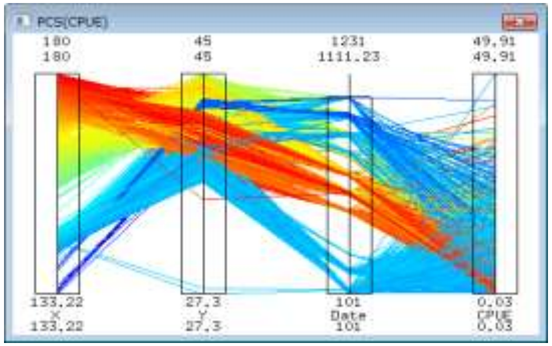


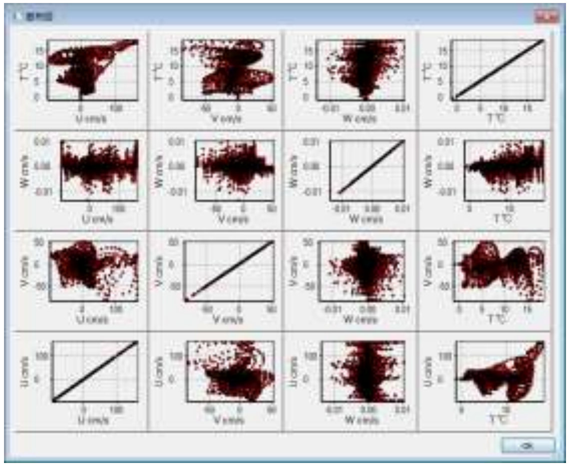


座標空間



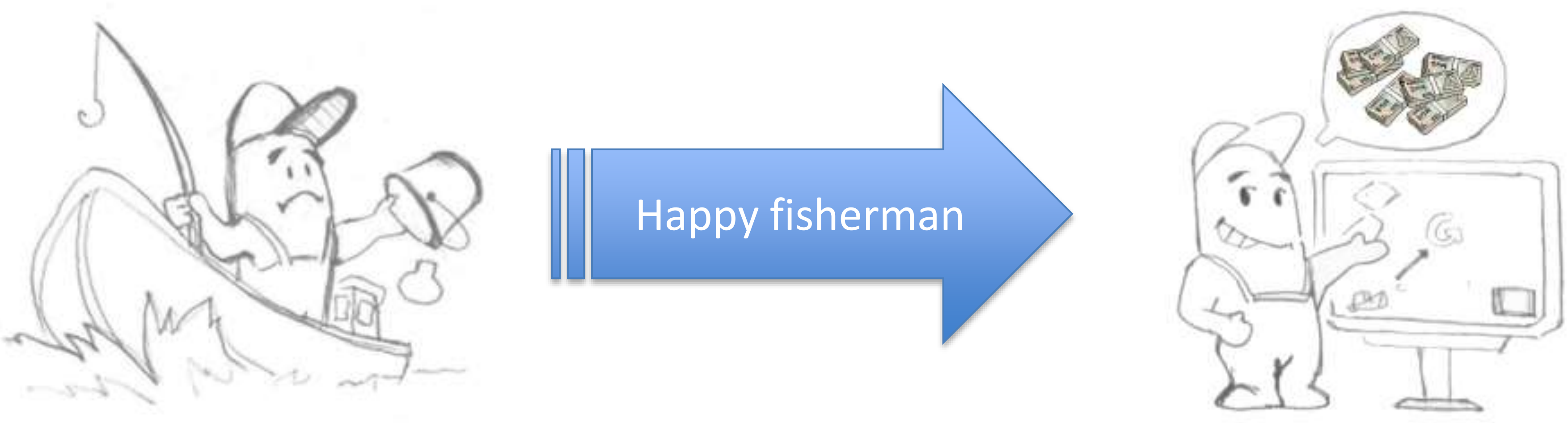
視覚的分析環境





変数空間

同化データ・計測データ



可視化の拓く世界

漁場予測への適用

HTTP://WWW.MEXT-ISACC.JP/_PUBLIC/EVENT2011_DATA/POSTER/RECCA_AWAJI.PDF

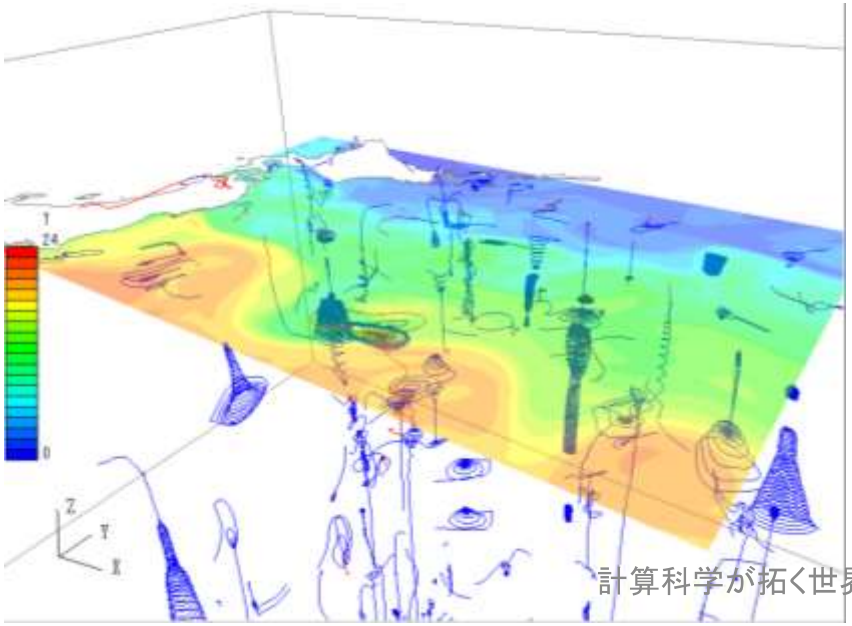
漁師の苦悩

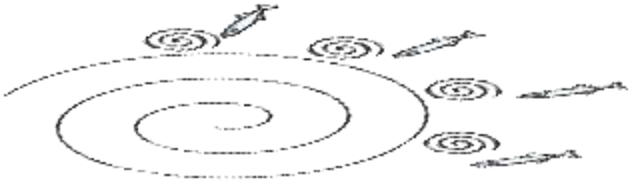
- 以前はこれまでの経験と勘で漁場に辿り着けた
- いまやそのような状況にはない

漁場探索技術によせる期待が大きい

漁師の勘と経験

- よい漁場は黒潮と親潮との潮目にできる渦の周辺に見つかる
- 地球温暖化の影響で潮目や渦の位置が以前と違ってきている



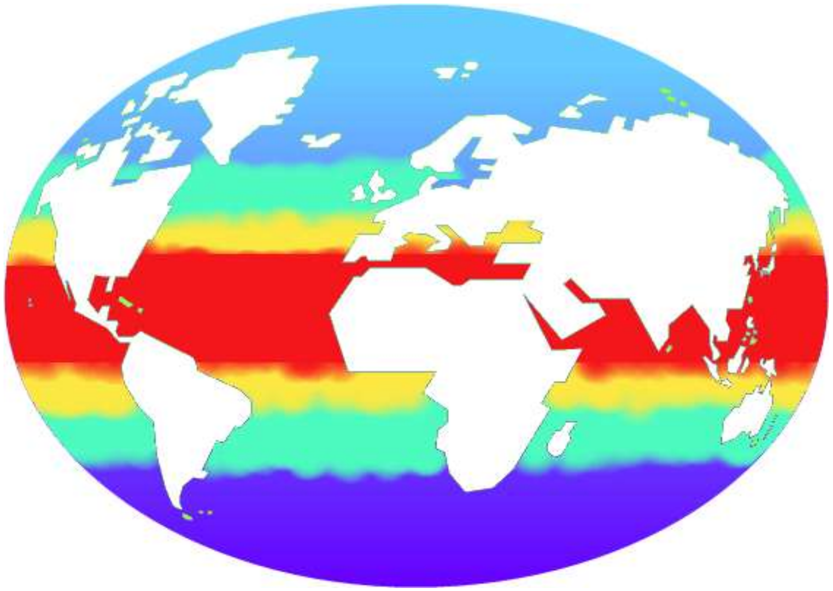


漁師の味方

- 漁場予測技術
  - 大気海洋シミュレーションモデルの精度向上
    - 先進モデリング技術
    - 高性能計算機環境

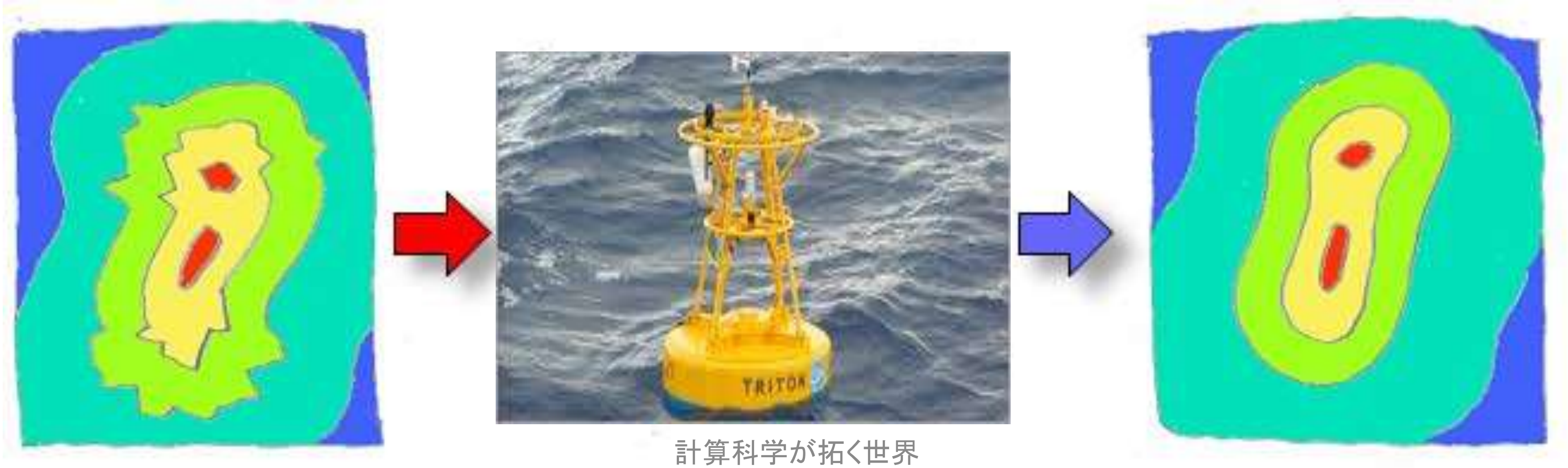
海洋モデリング技術

- モデル精度は関連データの計測結果を取り込むことにより向上させることができる。
- このような技術のひとつとしてデータ同化技術がある。



データ同化

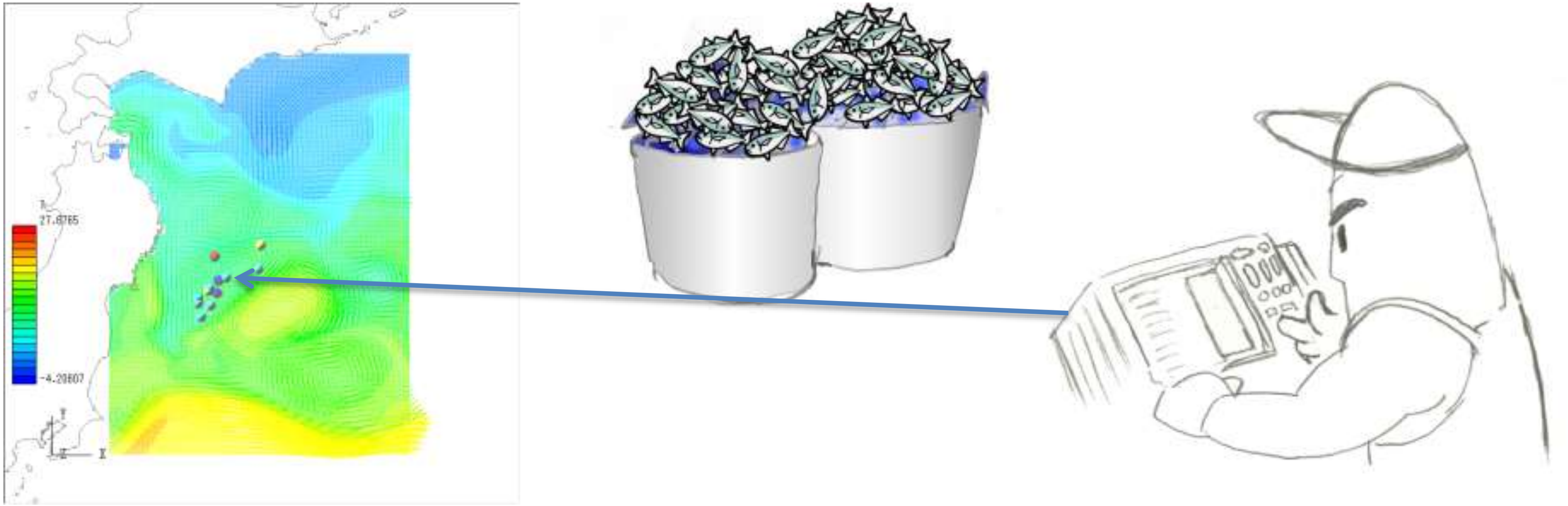
- データ同化は観察結果をコンピュータモデルに織りなす処理のことを指す。
- データ同化は、地球物理学における多くの分野で利用され、特に気象・海象予報で活用されている。



漁師による漁獲量計測

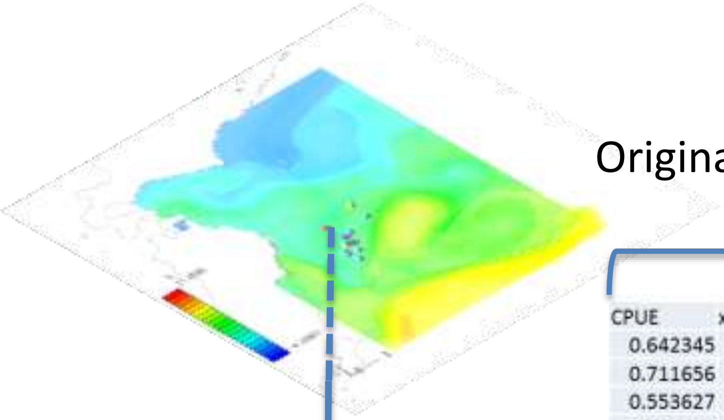
- 多くの漁師は以下の形式で漁獲量を計測し、記録
  - 漁船位置 (x,y)
  - 正規化漁獲量

$$CPUE = \frac{\sum catch}{\sum fishing\ \ days \sum fishing\ \ vessels}$$



CPUEをシミュレーション結果に含まれる変数で説明

- シミュレーション結果においてCPUEの定義される点を通る鉛直線上でサンプリング



(t_1, s_1, v_1)

$(t_{100}, s_{100}, v_{100})$

Original record　　Sampled variables

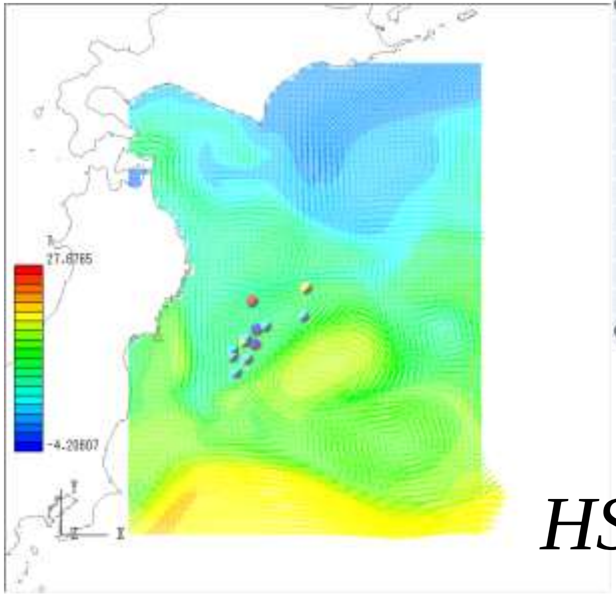
CPUE	x	y	t1	···	t100	s1	···	s100	v1	···	v100
0.642345	0.680518	0.755554	0.448537	···	0.920516	0.754839	···	0.113546	0.336512	···	0.214305
0.711656	0.283465	0.059334	0.099365	···	0.586174	0.890815	···	0.048899	0.536846	···	0.47709
0.553627	0.223513	0.395263	0.532903	···	0.985017	0.259018	···	0.884757	0.242159	···	0.699667
0.096016	0.086106	0.15024	0.131232	···	0.390091	0.958255	···	0.136579	0.506822	···	0.158313
0.927557	0.087157	0.576571	0.74707	···	0.311494	0.039833	···	0.265292	0.699515	···	0.433637
0.776234	0.66956	0.807943	0.680908	···	0.667134	0.189051	···	0.234158	0.83681	···	0.430659
0.339953	0.361467	0.427135	0.967647	···	0.251424	0.968415	···	0.437858	0.330161	···	0.401806
0.167646	0.512573	0.036514	0.060821	···	0.257233	0.103441	···	0.199014	0.339543	···	0.631877
0.525502	0.868313	0.127016	0.576511	···	0.665982	0.455188	···	0.379432	0.785053	···	0.58507
0.476963	0.512256	0.676056	0.301396	···	0.046038	0.187826	···	0.696615	0.44408	···	0.592365

$$CPUE = f\left(t_1, \cdots t_{100}, s_1, \cdots s_{100}, v_1, \cdots v_{100}\right)$$

総合好適度(HSI) モデル

Tian, S., Chen, X., Chen, Y., Xu, L., Dai, X. 2009. Fisheries Research

- HSI は正規化されたCPUEを表す
- CPUEと高い相関を持つ変数を選択



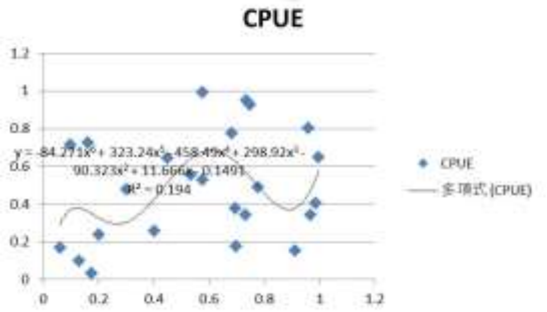
CPUE	x	y	t1	・・・	t100	s1	・・・	s100	v1	・・・	v100
0.642345	0.680518	0.755554	0.448537	・・・	0.920516	0.754839	・・・	0.113546	0.336512	・・・	0.214305
0.711656	0.283465	0.059334	0.099365	・・・	0.586174	0.890815	・・・	0.048899	0.536846	・・・	0.47709
0.553627	0.223513	0.395263	0.532903	・・・	0.985017	0.259018	・・・	0.884757	0.242159	・・・	0.699667
0.096016	0.086106	0.15024	0.131232	・・・	0.390091	0.958255	・・・	0.136579	0.506822	・・・	0.158313
0.927557	0.087157	0.576571	0.74707	・・・	0.311494	0.039833	・・・	0.265292	0.699515	・・・	0.433637
0.776234	0.66956	0.807943	0.680908	・・・	0.667134	0.189051	・・・	0.234158	0.83681	・・・	0.430659
0.339953	0.361467	0.427135	0.967647	・・・	0.251424	0.968415	・・・	0.437858	0.330161	・・・	0.401806
0.167646	0.512573	0.036514	0.060821	・・・	0.257233	0.103441	・・・	0.199014	0.339543	・・・	0.631877
0.525502	0.868313	0.127016	0.576511	・・・	0.665982	0.455188	・・・	0.379432	0.785053	・・・	0.58507
0.476963	0.512256	0.676056	0.301396	・・・	0.046038	0.187826	・・・	0.696615	0.44408	・・・	0.592365
Correlation			0.419613	・・・	0.34039	-0.33172	・・・	-0.03077	0.477163	・・・	0.08621

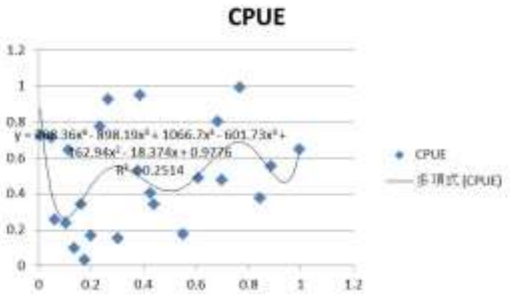
$$HSI = g(t_1, t_{100}, v_1) = \sqrt[3]{SI_1(t_1) SI_2(t_{100}) SI_3(v_1)}$$

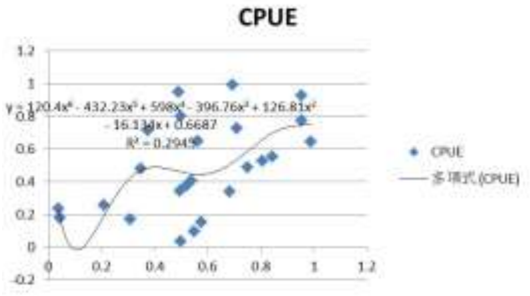
好適度(SI)

- SI 方程式はそれぞれの変数の回帰式で表現される。

$$HSI = \sqrt[3]{SI_1(t_1)SI_2(t_{100})SI_3(v_1)}$$







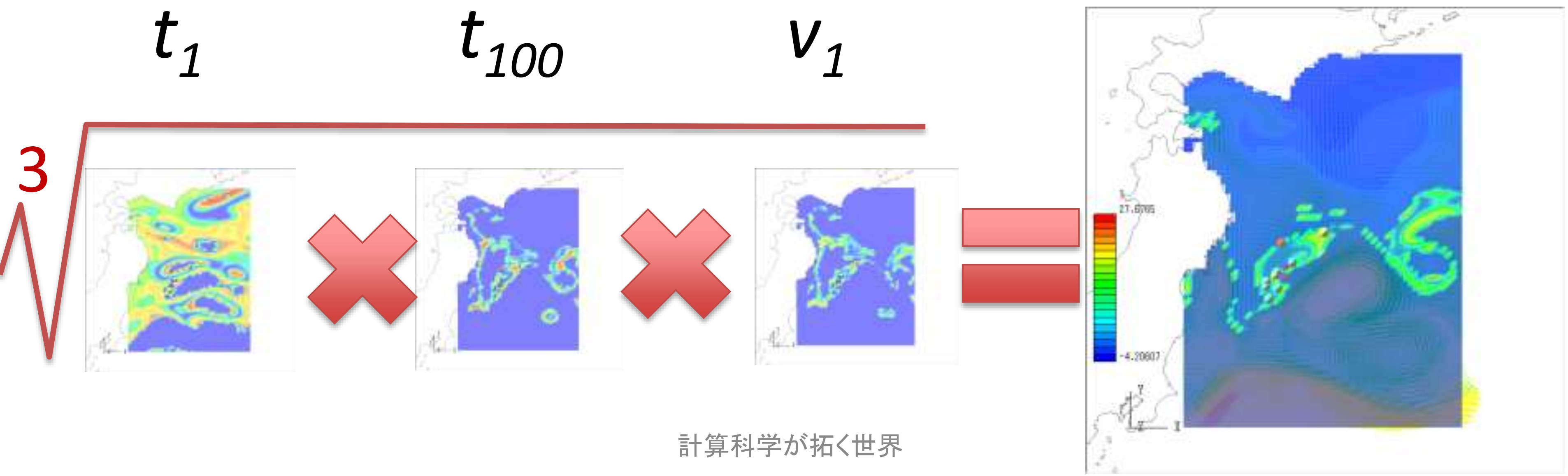
$$HSI = \sqrt[n]{\prod_{i=1}^{n} SI_i}$$

$$SI_i = \frac{CPUE_{fit_i} - \min CPUE_{fit_i}}{\max CPUE_{fit_i} - \min CPUE_{fit_i}}$$

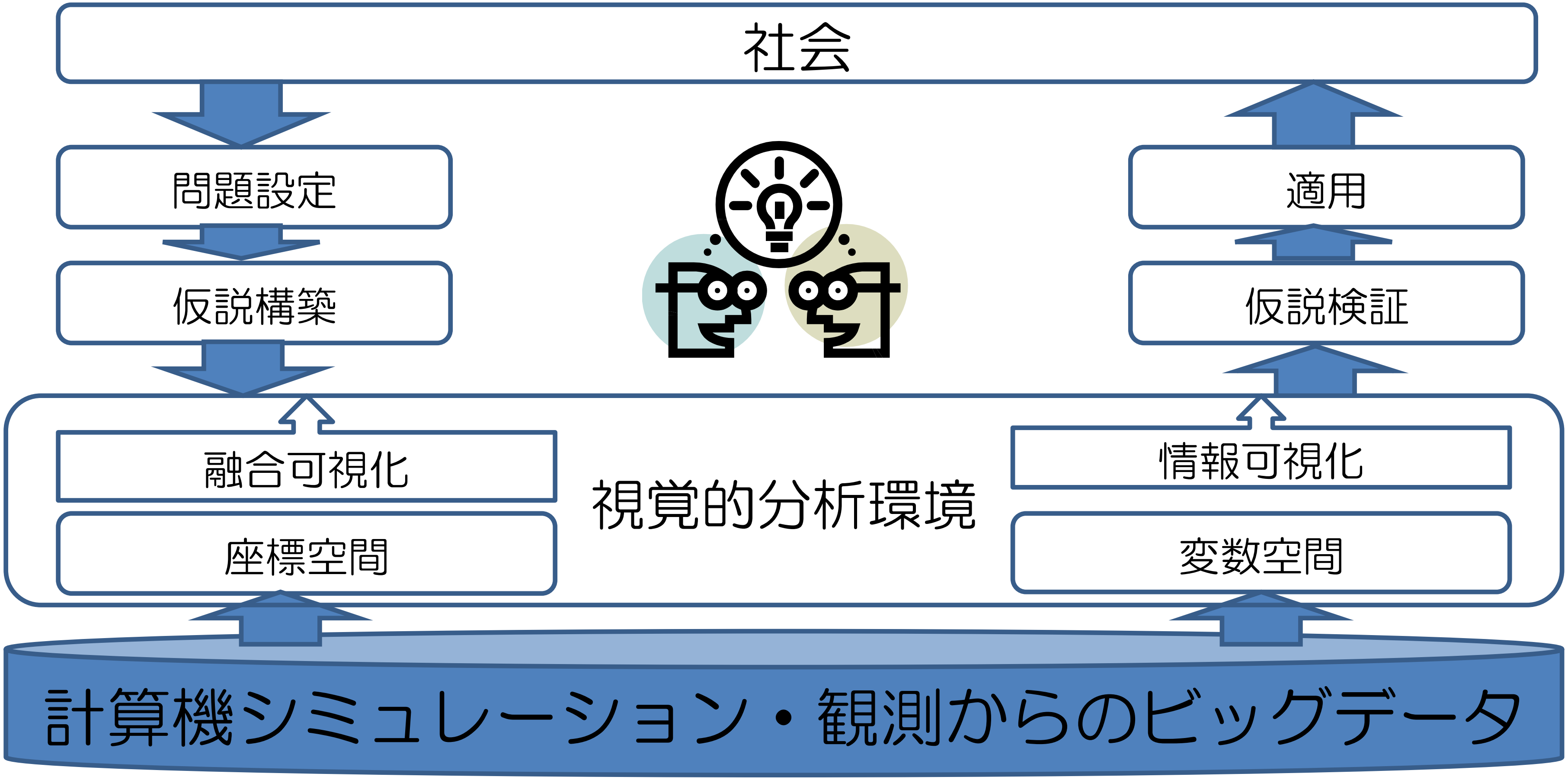
HSI モデル開発の手順

以下の手順でモデルを開発する:

1. CPUEと高い相関を持つ変数を選択し、SI方程式を構築
2. SI方程式の相乗平均としてHISを構築
3. モデル妥当性検証のためにHIS分布を可視化



仮説検証支援用視覚的分析環境



まとめ

- 可視化技術の基礎と応用について理解した
  – 可視化について
  – 可視化事例
  – 視覚的分析環境
- 興味があれば、計算科学シミュレーション結果の社会実装（たとえば、科学技術イノベーション政策など）を考えてみよ。
  – 参考：「社会実装」という言葉は、科学技術振興機構の「社会技術」という概念から生まれてきた。「社会技術」とは、「人間や社会のための科学技術」という意味である。