

共分散構造分析は 自然科学からモノづくりへ

筑波大学大学院ビジネス科学研究科
国際経営プロフェッショナル専攻長
(併)統計数理研究所リスク解析戦略研究センター長
椿 広計

2006/10/25 JUSE-StatWorks/V4.0 SEM
因果分析編製品発表説明会において

祝辞

- 私は、特定の統計ソフトにくみしませんが、とにかく、因果分析の二大潮流である、共分散構造分析(SEM:表線形)とグラフィカルモデル(GM:裏線形)とが合体した世界初の応用向けソフトが出来たことは、私にとっても念願がかなったことで大変うれしく思います。他の統計ソフトウェアも、ぜひ新たな考え方を導入し、SEMとGMとの結合という未開分野に進出して頂ければと思います。
- 今回のソフトウェアは、共分散構造分析分野ではAMOSと並ぶ2大エンジンであるBentler博士のEQSと、(社)日本品質管理学会テクノメトリックス研究会、特に東京工業大学の宮川雅己教授の理論的支援を受け、芳賀敏郎先生が開発したCGGMを基に、広野元久氏が種々の考え方を追加したGGMを結合したツールです。この両者が、科学技術分野で活用されるまでには、少し時間がかかるかもしれませんが、技術者が創造的に考えるツールとしての新SQCに育つのではないかと、大変楽しみにしています。
- 私がなぜこのような考えを持つに至ったか、今日は、この分野に関する少し個人的な思い出話に付き合って頂ければと思います。

私の現在の考え

因果分析のニーズ：回帰分析の問題

- 回帰分析でデータから導かれる「予測式」と技術者が想像している現象のメカニズムにはギャップが生じている
 - 回帰分析で用いる変量間の「因果関係 (Causality)」の無視
 - 「測定結果 (Measurement Results)」と「真値 (True Value)」との混同
- 「科学的立場のデータ解析 (Explanatory Approach)」の必要性

構造方程式モデリング

- 因果分析・共分散構造分析・
構造方程式モデリングの特徴
- 回帰分析との比較
 - 全体的な因果関係の理解が重要
 - 単一方程式→連立方程式
 - 測定モデルの導入

本日の話題

- 計測工学で学んだこと
 - 説明変数に誤差のある回帰分析
 - 因子分析と回帰分析
 - なぜ因果分析が必要なのか
- 再びリモートセンシングへ
 - 未知物質のモデル化に挑む
 - 本当に因子分析になってしまった自然科学
- 結語：モノづくりへの適用の期待

統計の大学院は出たものの —計測工学で学んだこと—

- 1982年3月
 - 東京大学工学部計数工学科奥野忠一教授,
廣津千尋助教授(統計工学)研究室修士課程修了
 - タグチメソッドの統計的側面
 - 新医薬品の許認可に使われている検定
 - 博士課程の受験に失敗:統計学を断念
- 1982年4月
 - 計数工学科計測工学通論講座(豊田弘道教授, 藤村貞夫助教授)助手に拾って頂く
 - 温度計測+遠隔環境計測(リモートセンシング)の研究
 - 学会全体に, 多変量解析のパワーユーザーが揃っている事に驚く
 - 8月:第4回RESESで回帰診断に関するチュートリアル「統計学夜話」

単回帰分析における 説明変数と目的変数

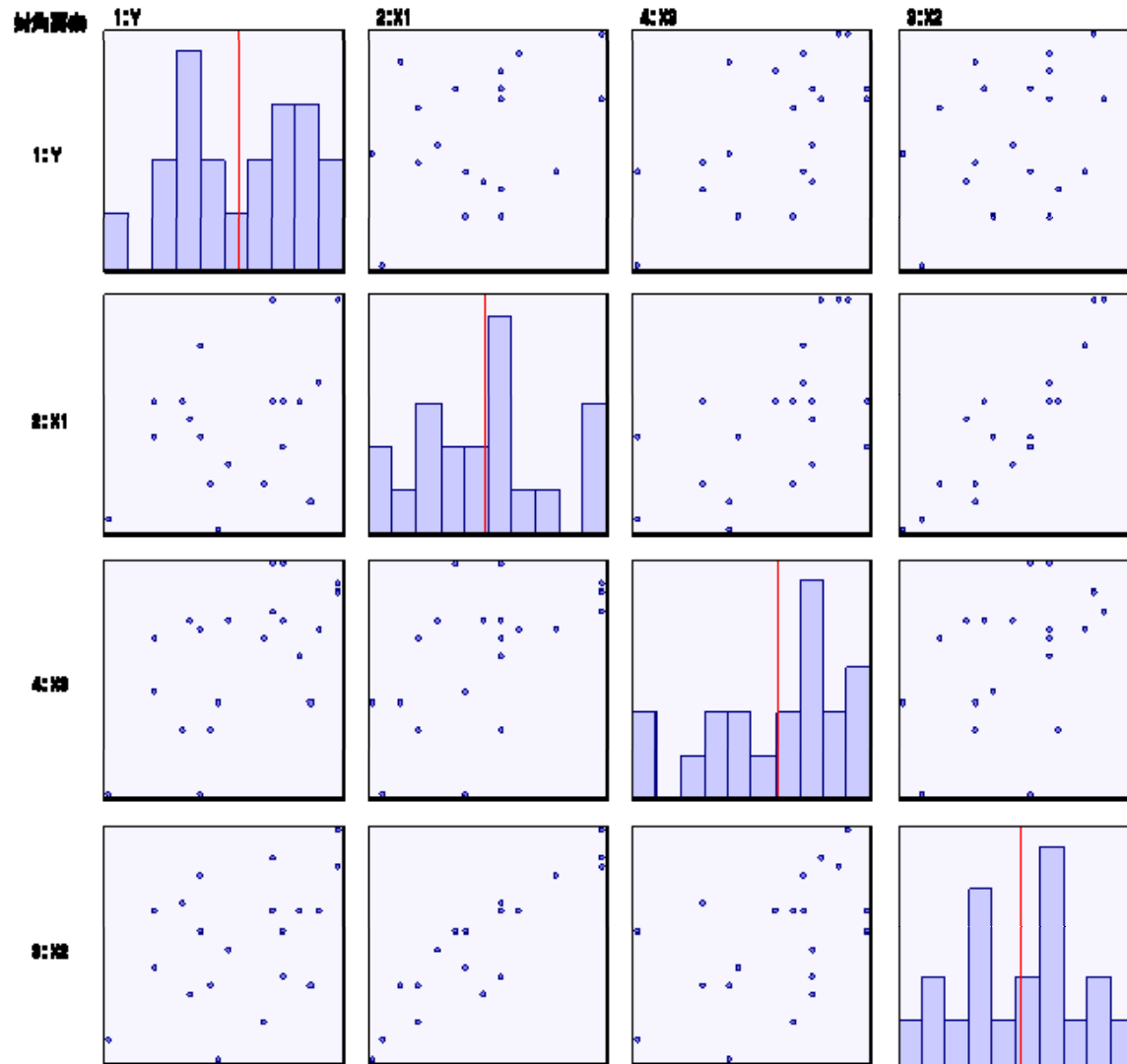
奥野先生の主張

回帰分析における 説明変数と目的変数

- 1982年11月：計測自動制御学会第8回リモートセンシングシンポジウム
奥野忠一先生「リモートセンシングデータの統計解析における諸問題」特別講演に招待
 - 椿, 奥野, 安岡(1983)「リモートセンシングにおけるデータ解析—回帰分析の適用を中心にして」, 日本リモートセンシング学会誌第3巻, pp. 31–43を執筆
 - » リモートセンシング学会10周年時に表彰対象
- リモートセンシングにおける回帰分析の目的変数は、計測したい実情報 (Ground-Truth) か、衛星推定情報 (分光情報) か？
 - 奥野先生の回答：衛星推定情報
 - 一般的なQC教育：予測したいものを目的変数にしなさい

因果関係考察の無視がもたらすもの

炉内温度計測問題: Y, 炉内温度, X1~X3炉外温度



重回帰分析の適用

藤村貞夫先生：統計はそんな考え方なの？

重回帰分析の単純な適用結果

偏回帰係数の推定結果：

	推定値	標準誤差	t 値	p 値
(Intercept)	120.82462	55.20212	2.189	0.0438 *
X1	-0.05431	0.43253	-0.126	0.9016
X2	0.56450	0.25792	2.189	0.0438 *
X3	0.18027	0.36516	0.494	0.6282

残差標準偏差： 22.73 (自由度 16)

寄与率： 0.3906, 自由度調整済み寄与率： 0.2763

F-統計量： 3.418 分子の自由度 3 分母の自由度 16 , p-値： 0.049

変数選択→単回帰分析

偏回帰係数の推定結果:

	推定値	標準誤差	t 値	p 値
(Intercept)	131.4809	49.6359	2.649	0.01633 *
X2	0.6476	0.2010	3.222	0.00473 **

残差標準偏差: 21.86 (自由度 18)

寄与率: 0.3657, 自由度調整済み寄与率: 0.3307

F-統計量: 10.38 分子の自由度 1 分母の自由度 18 , p-値: 0.0004

因果関係を尊重した単回帰分析

奥野先生の主張：目的変数→結果変数

予測は逆回帰（回帰式を逆に解く）で行う：結果は大違い！

偏回帰係数の推定結果：

	推定値	標準誤差	t 値	p 値
(Intercept)	81.6045	51.1476	1.595	0.12801
Y	0.5647	0.1753	3.222	0.00473 **

残差標準偏差：20.42（自由度 18）

寄与率：0.3657， 自由度調整済み寄与率：0.3307

F-統計量：10.38 分子の自由度 1 分母の自由度 18， p-値：0.0004

説明変数にも測定誤差はある

知らしめてはならない問題

藤村先生と横山先生との論争

藤村理論による予想と実証実験

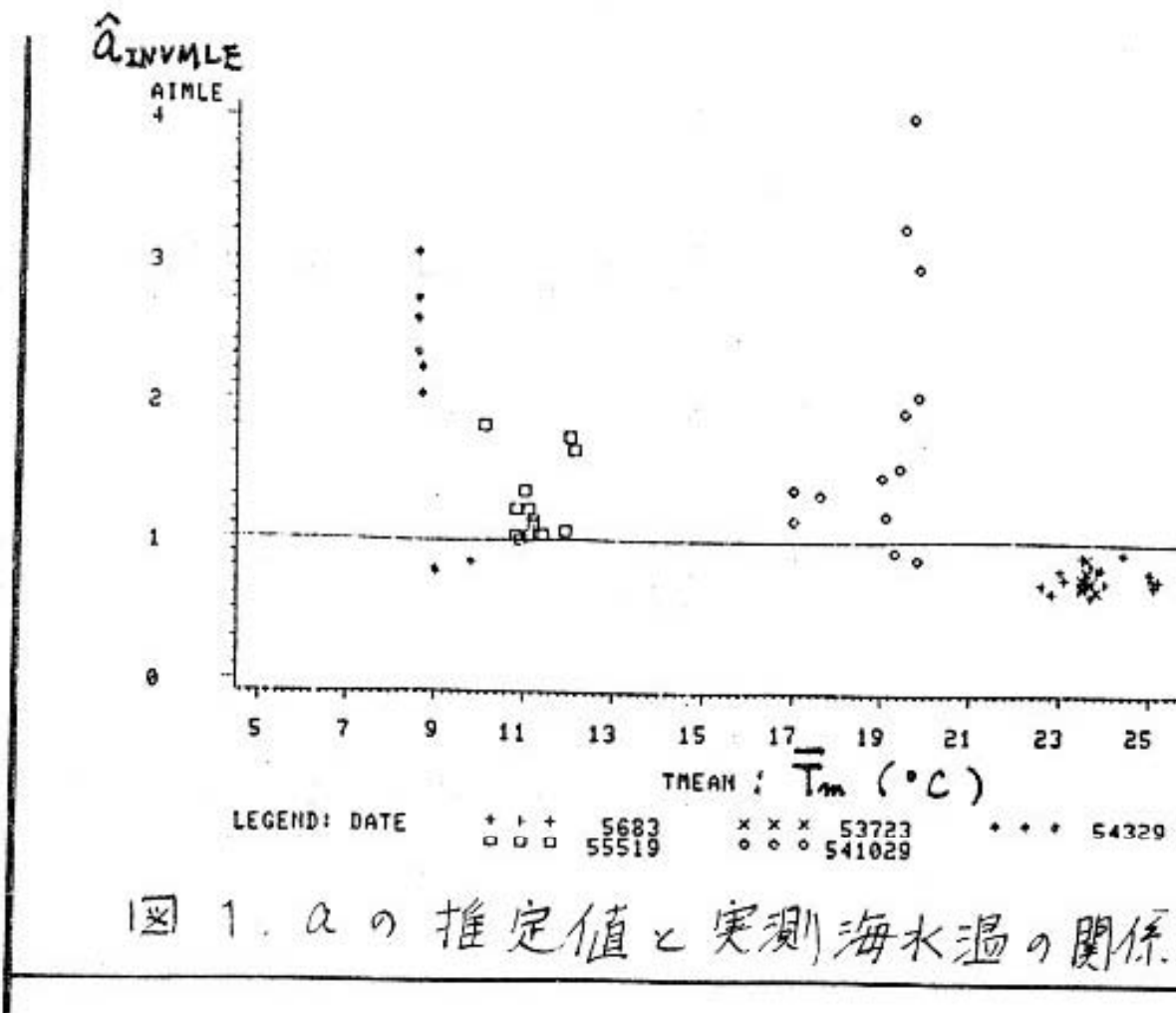
- 海面上のある地点の温度を測定する方法
 - バケツ水温, 海上実測水温: t_m
 - その地点の水をバケツで汲み上げ温度計で測定
 - 見かけの水温, 推定海面温度: t_e
 - 航空機から撮影した熱赤外画像
- 藤村(1983)の応用物理論: $t_e = (\varepsilon\tau)^{1/\beta} t_m$
 - 第9回計測自動制御学会リモートセンシングシンポジウム
 - » 藤村,横田(1986)熱赤外遠隔計測データからの温度推定モデル, 計測自動制御学会論文集, 22巻, 779-784.
 - ε : 水面物体放射率, τ : 大気透過率
 - β : 熱放射拡張指数則の係数
 - 椿, 花泉, 藤村, 豊田(1982)第8回リモートセンシングシンポジウム
 - 藤村, 花泉, 椿(1985)熱放射の拡張指数則
計測自動制御学会論文集, 21巻, 208-210
 - 樋富, 藤村他(1982)の水槽実験, 第4回RESESシンポジウム
 - 実験データの回帰分析: $T_e = \alpha_0 + \alpha_1 T_m + \varepsilon_{Te}$
 - 回帰係数 α_1 の推定値は1未満を実証
 - » 証明, 風, 攪拌, 気温・水温差といった誤差因子を導入
→回帰係数を1未満にする方向に作用

横山隆三先生の観測との矛盾

- 横山(1982)
 - 岩手大学教授
 - 第2回日本リモートセンシング学会講演
 - 三陸沿岸で62回の航空機リモートセンシング観測
 - 観測データセット毎に回帰分析を62回行う
 - $T_m = \gamma_0 + \gamma_1 T_e + \varepsilon_{Tm}$
 - » $(\varepsilon\tau)^{1/\beta} \sim 1/\gamma_1$
 - 異常値の排除など回帰診断も実施
 - 季節, 観測場所, 観測時刻, 観測高度などで層別し結果報告
 - 冬季を中心に, この値は1より大きくなる傾向がある
 - 藤村理論との矛盾
 - 藤村先生の反発
 - » そんな馬鹿なことはいえない!
 - » 回帰分析の説明変数と目的変数の取り違えが問題?
 - データを借りるために藤村先生と岩手出張
 - » 奥野流に回帰分析の説明変数と目的変数を変えても横山説は再現!

横山先生が示した散布図

椿, 花泉, 藤村(1984)応用統計学会年会予稿で
再計算作図



説明変数に誤差のある回帰分析 問題は統計モデルに！

- 修士時代の輪講を思い出す！
 - Kendall and Stuart (1981) *The Advanced Theory of Statistics*, Vol.2 4th ed., Griffin.
 - 回帰モデルは、関数関係モデル、あるいは構造関係モデルの説明変数の測定誤差を無視したモデル
 - この問題を最初期に指摘したのは、次と思われる：
Deming(1943) *Statistical Adjustment of Data*, Wiley.

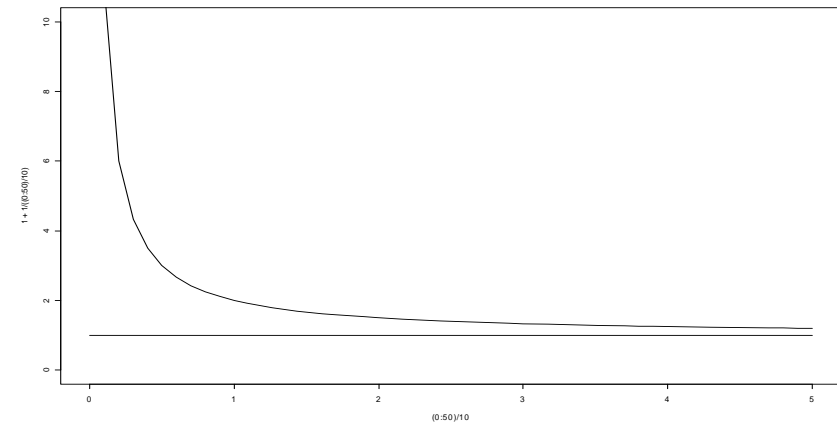
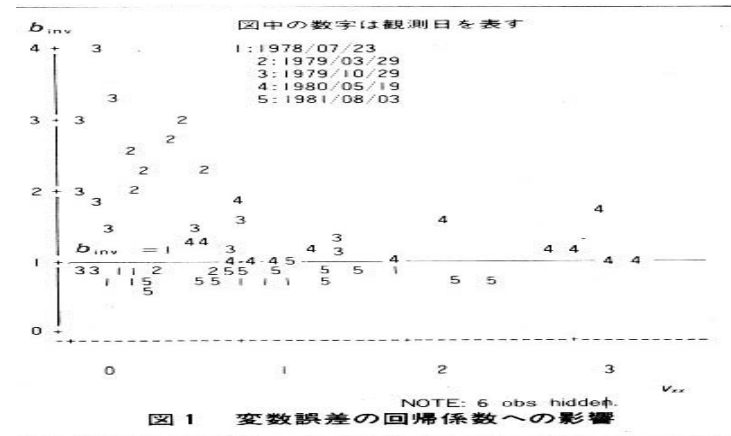
構造関係モデル

説明変数に誤差のある回帰分析

- 通常の最小二乗法では回帰係数の過小評価

- $T_m = t_m + \varepsilon_{Tm}, \varepsilon_{Tm} \sim N(0, \sigma_{Tm}^2)$
- $T_e = t_e + \varepsilon_{Te}, \varepsilon_{Te} \sim N(0, \sigma_{Te}^2)$
- $t_e = \alpha_0 + \alpha_1 t_m, t_m \sim N(\mu_{tm}, \eta_{tm}^2)$
 - SN比: $\lambda_m = \eta_{tm}^2 / \sigma_{Tm}^2$

- 回帰係数の不一致性:
 $a_1 = S_{TmTe} / S_{TmTm} \rightarrow \lambda_m \alpha_1 / (1 + \lambda_m)$
- 回帰係数の逆数は, $(1 + 1/\lambda_m) \alpha_1^{-1}$ となり過大評価
 - 藤村先生の散布図
 - 水面温度の分散が小さいほど回帰係数の逆数が大きくなる



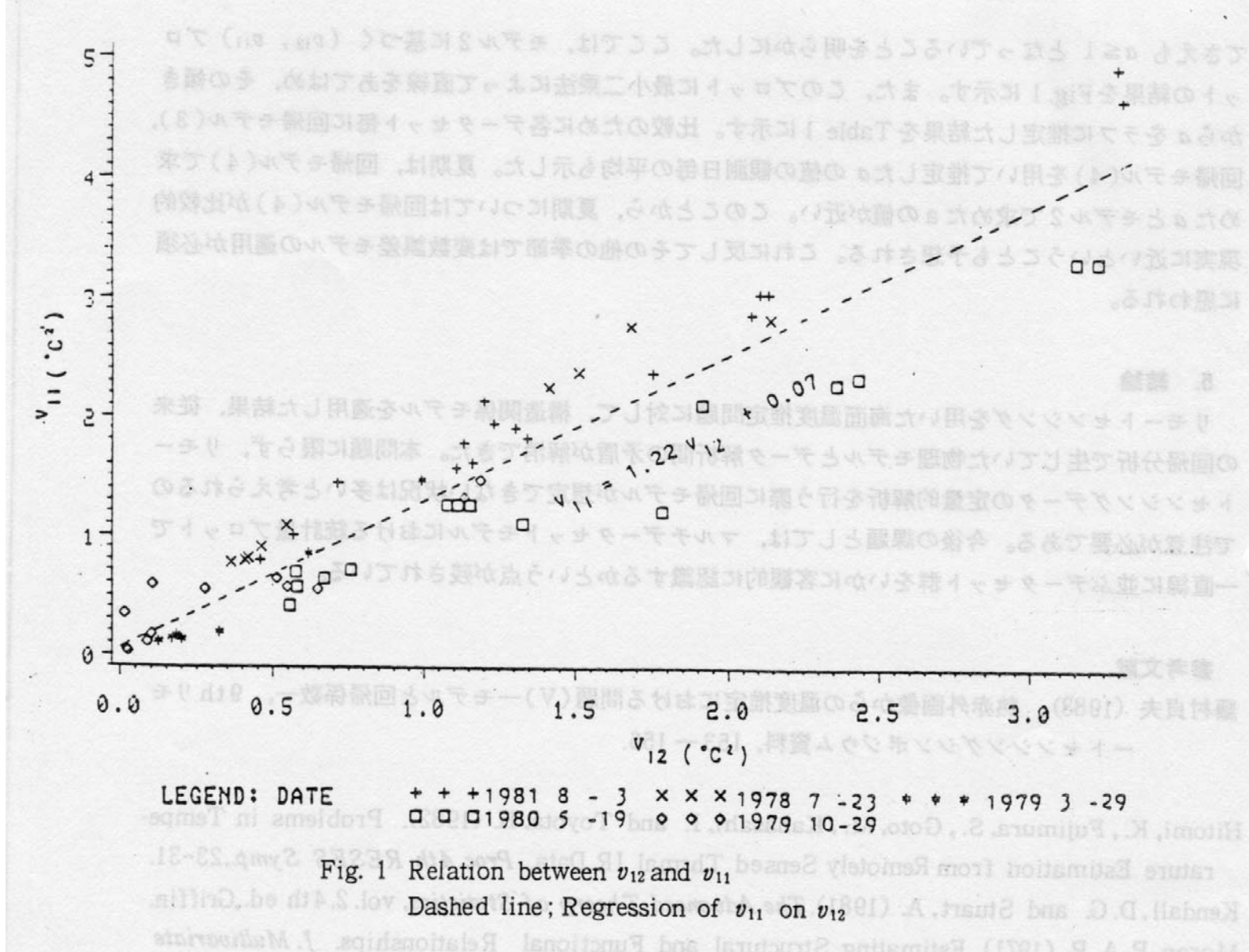
その後2回この問題を見かける

- 日次電力使用量と気温との関係
 - (東京ガス・樫尾さんの日本規格協会データ解析研究委員会事例)
 - 6-8月の3ヶ月データの温度と電力使用量の温度に対する回帰係数は、6月、7月、8月の各月毎に求めた3つの回帰係数のいずれよりも大きい
- 日銀金融研究所の研究會
 - 説明変数: 金利変動をm日間のデータの標準偏差で表現
 - mを大きくすると回帰係数は大きくなる

説明変数に誤差のある単回帰分析は解けない 本当の回帰係数はSN比依存

- 解けないものを解く仮定
 - 2つのマルチデータセットモデルの提案
 - 構造: $t_e = \alpha_0 + \alpha_1 t_m$ が幾つかのデータセットで共通
 - α_1 とバケツ温度の誤差分散 σ_{Tm}^2 が幾つかのデータセットで共通
 - 今日でいう多母集団分析
 - 椿, 横山, 花泉, 藤村, 豊田(1983)リモートセンシングデータを用いた海面温度推定のための統計モデル, 第9回リモートセンシングシンポジウム, 157-158
 - 椿, 花泉, 藤村(1984)グランドトゥースとリモートセンシングデータを結ぶ統計モデル, 第10回リモートセンシングシンポジウム招待講演, 175-178.
 - 当時は全データの最尤法ではなく
2データセット毎の最尤法に基づく診断プロットを提唱

実測・推定共分散VS実測値分散プロット



プロットから読み取った傾向

Table 1

Comparison of estimates different statistical model
using data from Yokoyama (1982)

Table 1

Comparison of estimates of a from different statistical model
using data from Yokoyama(1982)

Date	No. of Datasets	Estimates of a		
		from Model 2 (18)	Mean of \hat{a} from Model (3)	Mean of \hat{a} from Model (4)
1978-7-23	9	0.769	0.572	0.744
1979-3-29	8	0.613	1.261	2.065
1979-10-29	13	0.971	0.899	1.845
1980-5-19	16	0.990	1.044	1.221
1981-8- 3	16	0.730	0.665	0.772
pooled	62	0.820	0.875	1.276

診断プロット上の
回帰分析
藤村理論と整合的

因果関係は
正しい
回帰モデル

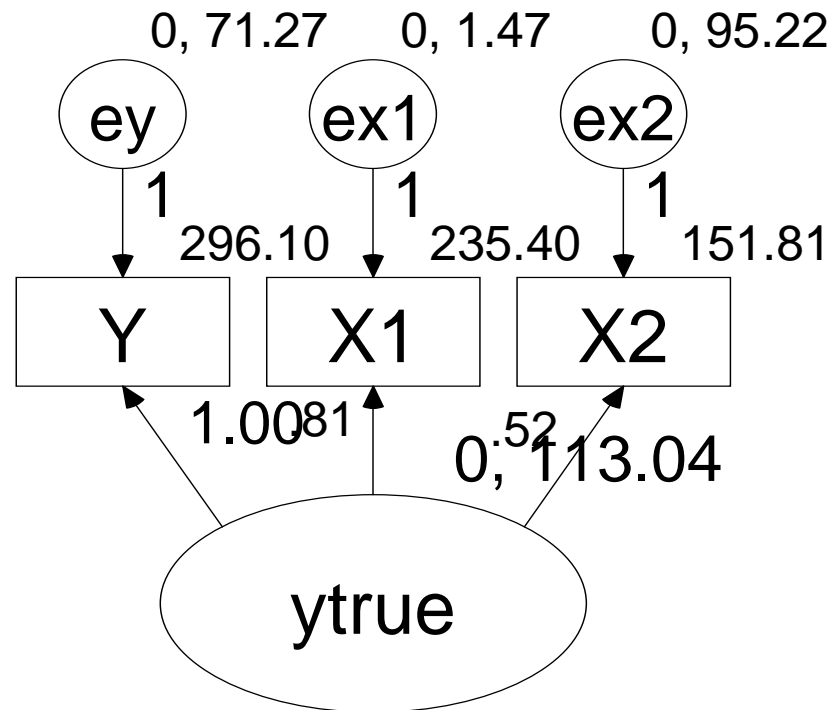
技術的目的に
即した
回帰モデル

シングルデータセットでは？
「操作変数法」
マルチスペクトル変数誤差モデル

計測工学：マルチスペクトラルモデル

計量経済：操作変数法

計量心理：検証的因子分析



「測定方程式」 真値が因子

- $X1 = \text{factor} + e1$
- $X2 = b0 + b1 \times \text{factor} + e2$
- $X3 = c0 + c1 \times \text{factor} + e3$
 - 「基準変量(Criterion Variable)」
 - $X1$ は潜在変数自体を誤差 $e1$ を伴って、直接計測した変量と考える
 - 潜在変数（真値）は、実際には観測されていない
 - その「単位」あるいは平均や分散には不定性
 - この不定性を回避するために、何らかの観測変数（バケツ温度のような変数）を基準変量として導入

炉内温度推定モデルの結果 (最尤法)

:

	推定値	標準誤差	検定統計量	確率
Y ←--- Y の真値	1.000			
X1 ←--- Y の真値	1.108	.449	2.468	.014
X2 ←--- Y の真値	1.198	.495	2.421	.015

平均値 :

	推定値	標準誤差	検定統計量	確率
Y の真値	290.612	5.975	48.640	***

切片 :

	推定値	標準誤差	検定統計量	確率	ラベル
X1	-52.774	130.630	-.404	.686	
X2	-102.339	143.904	-.711	.477	

分散:

	推定値	標準誤差	検定統計量	確率
Y の真値	319.810	215.657	1.483	.138
eY	358.438	158.533	2.261	.024
eX1	509.786	211.718	2.408	.016
eX2	132.712	160.443	.827	.408

因子分析との関係

- 構造関係(変数誤差)モデルを解くために計量経済学で用いられている操作変数法は, 因子分析そのものと確信
- 変数誤差モデルは実在する因子分析?
 - 行動計量学会年次大会1986年夏東大で開催
 - ナイトセッション
 - 因子は実在するのですかフィクションなのですか?
 - » 計量心理の先生方がフィクションと応えられ大いに失望
- 我が意を得たり
 - T.W. Anderson (1988) Wald Memorial Lecture, Annals of Statistics:
因子分析と構造関係モデルは統計モデルとして同等という前提で議論

横山先生との和解と反論(1984)

椿さん, 理屈は分かりました.
でも予測式は技術的重回帰の変数選択で
十分でしょう. なぜならば
回帰予測は最良予測なのです!

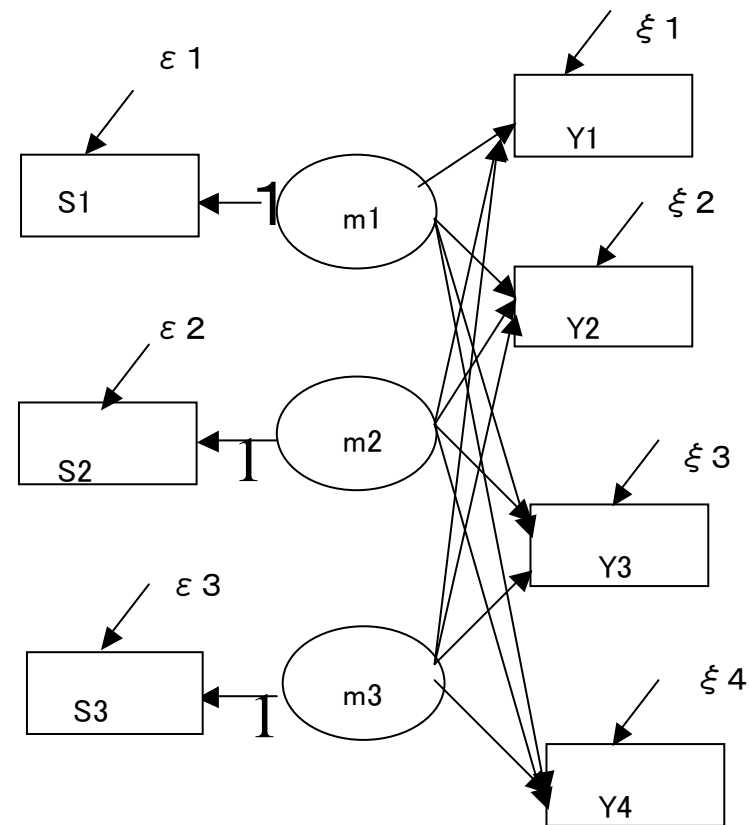
これには参りました

なぜ因果分析は技術にとって有用なのか？ なぜ、リモートセンシングの回帰予測式には 再現性がないのか？

回帰予測か，因子得点か？：逆問題で考えたこと
椿(1998)リモートセンシングにおける統計モデル(1)
統計手法から統計モデリングへ，
日本リモートセンシング学会誌，18巻1号，57～64.
椿(2000)「逆問題における因果推論の意義」
統計学会予稿集

簡明な因果モデルからの出発 計測行為を意識して

- q 入力 p 出力Latent Variables Model
- 入力： $\mathbf{S}=\mathbf{m}+\boldsymbol{\varepsilon}$
- 出力： $\mathbf{Y}=\mathbf{B}\mathbf{m}+\boldsymbol{\alpha}+\boldsymbol{\xi}$
- $\mathbf{m} \sim N_q(\boldsymbol{\mu}, \boldsymbol{\Omega})$
- $\boldsymbol{\varepsilon} \sim N_q(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon)$
- $\boldsymbol{\xi} \sim N_p(\mathbf{0}, \boldsymbol{\Sigma}_\xi)$
- $\boldsymbol{\varepsilon} \perp \boldsymbol{\xi}$



計測技術は典型的な逆問題

- ・ 計測問題

- 出力 Y が与えられたとき,
入力 m の値を予測する方式の確立
 - ・ 単純に考えれば： $E[m|Y]$ の推定問題
 - 横山先生の仰った通り回帰予測を行えば良いのか？
- $m \sim N_q(\mu, \Omega)$
 - ・ μ, Ω といった母数 (標示母数)
→ トレーニングデータを採ったときの環境に依存
- 計測 (予測) 式の一般化可能性
 - ・ データ取得環境に依存しない予測方式が望ましい

一般化可能性と最良回帰予測

- $E[m|Y] = (\Omega^{-1} + B^T \Sigma_{\xi}^{-1} B)^{-1} \Omega^{-1} \mu + (\Omega^{-1} + B^T \Sigma_{\xi}^{-1} B)^{-1} B^T \Sigma_{\xi}^{-1} (Y - \alpha)$
- $Cov[m|Y] = (\Omega^{-1} + B^T \Sigma_{\xi}^{-1} B)^{-1}$
 - これらの予測関数にトレーニング標本から推定した母数の最尤推定を代入すれば、単純な多変量回帰予測に帰着する
 - 多変量回帰予測は因果モデリングを省略可能？
 - 多変量回帰予測が不偏なのは、予測を行う場での入力信号の分布が、トレーニングデータの分布と等しい場合！

幾つかの考え方

- Copas: 予測を行う場のパラメータを用いる
 - 重要な指針
- 予測を行う場に依存しない方式を用いる
 - どのようにして達成？
 - 標示母数 // Ω^{-1} // $\rightarrow 0$: 無情報的予測状況
 - $E[m|Y] = (\mathbf{B}^T \Sigma_{\xi}^{-1} \mathbf{B})^{-1} \mathbf{B}^T \Sigma_{\xi}^{-1} (\mathbf{Y} - \alpha)$
 - $\text{Cov}[m|Y] = (\mathbf{B}^T \Sigma_{\xi}^{-1} \mathbf{B})^{-1}$
 - 予測誤差共分散行列は、予測の場の情報を使わない分だけ丁度大きくなる
 - 標示母数の値に依らず不偏予測である
 - 単純な回帰予測ではないので、因果モデルをどのように行ったかが本質的になる
 - » 奥野先生の考え方の拡張

m の \mathbf{S}, \mathbf{Y} を与えた下での 条件付期待値, 共分散行列

- 単純回帰予測

- $E[m|\mathbf{S}, \mathbf{Y}] =$
 $(\mathbf{\Omega}^{-1} + \mathbf{\Sigma}_{\varepsilon}^{-1} + \mathbf{B}^T \mathbf{\Sigma}_{\xi}^{-1} \mathbf{B})^{-1} \mathbf{\Omega}^{-1} \boldsymbol{\mu} +$
 $(\mathbf{\Omega}^{-1} + \mathbf{\Sigma}_{\varepsilon}^{-1} + \mathbf{B}^T \mathbf{\Sigma}_{\xi}^{-1} \mathbf{B})^{-1} \{ \mathbf{\Sigma}_{\varepsilon}^{-1} \mathbf{S} + \mathbf{B}^T \mathbf{\Sigma}_{\xi}^{-1} (\mathbf{Y} - \boldsymbol{\alpha}) \}$
- $\text{Cov}[m|\mathbf{Y}] = (\mathbf{\Omega}^{-1} + \mathbf{\Sigma}_{\varepsilon}^{-1} + \mathbf{B}^T \mathbf{\Sigma}_{\xi}^{-1} \mathbf{B})^{-1}$

- 標示母数に依存しない予測

- $E[m|\mathbf{S}, \mathbf{Y}] =$
 $(\mathbf{\Sigma}_{\varepsilon}^{-1} + \mathbf{B}^T \mathbf{\Sigma}_{\xi}^{-1} \mathbf{B})^{-1} \{ \mathbf{\Sigma}_{\varepsilon}^{-1} \mathbf{S} + \mathbf{B}^T \mathbf{\Sigma}_{\xi}^{-1} (\mathbf{Y} - \boldsymbol{\alpha}) \}$
- $\text{Cov}[m|\mathbf{Y}] = (\mathbf{\Sigma}_{\varepsilon}^{-1} + \mathbf{B}^T \mathbf{\Sigma}_{\xi}^{-1} \mathbf{B})^{-1}$

幾つかの留意点

- 因子得点推定の問題との類似性
 - ◎ Bartlettの方法(母数的因子)
 - × Thomsonの方法(変量的因子)
- m を変量ではなくて, 所与の定数と見なす
 - 攪乱母数の影響を排除するため?
- 順問題 $E[Y|m]$ では, 回帰関数が攪乱母数に元々依存しない.

琵琶湖水質環境の遠隔計測への 適用(椿, 1998)

- 安岡善文、宮崎忠国(1982)リモートセンシングによる湖水域の水質分布計測—回帰分析を用いた大気補正と水質推定、日本リモートセンシング学会誌、2巻3号、51-62.
- ランドサット衛星の多重分光画像データ
 - バンド4 : 0.5 ~ 0.6 μm ,
 - バンド5 : 0.6 ~ 0.7 μm ,
 - バンド6 : 0.7 ~ 0.8 μm)
- 琵琶湖水質データ
 - 琵琶湖の透明度 (cm)
 - SS (浮遊懸濁物質量:mg/l)
 - クロロフィル a 量

散布図行列

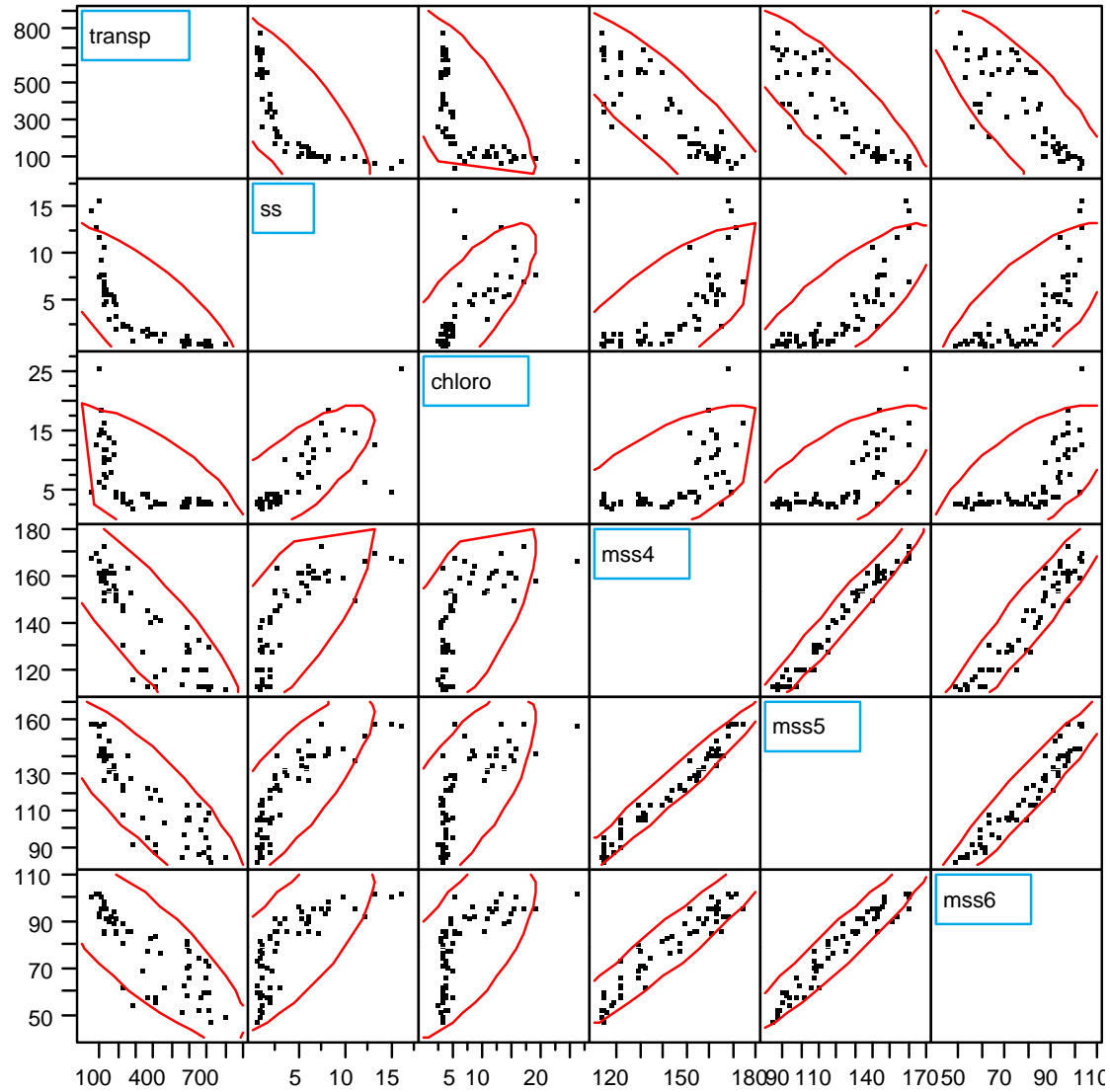
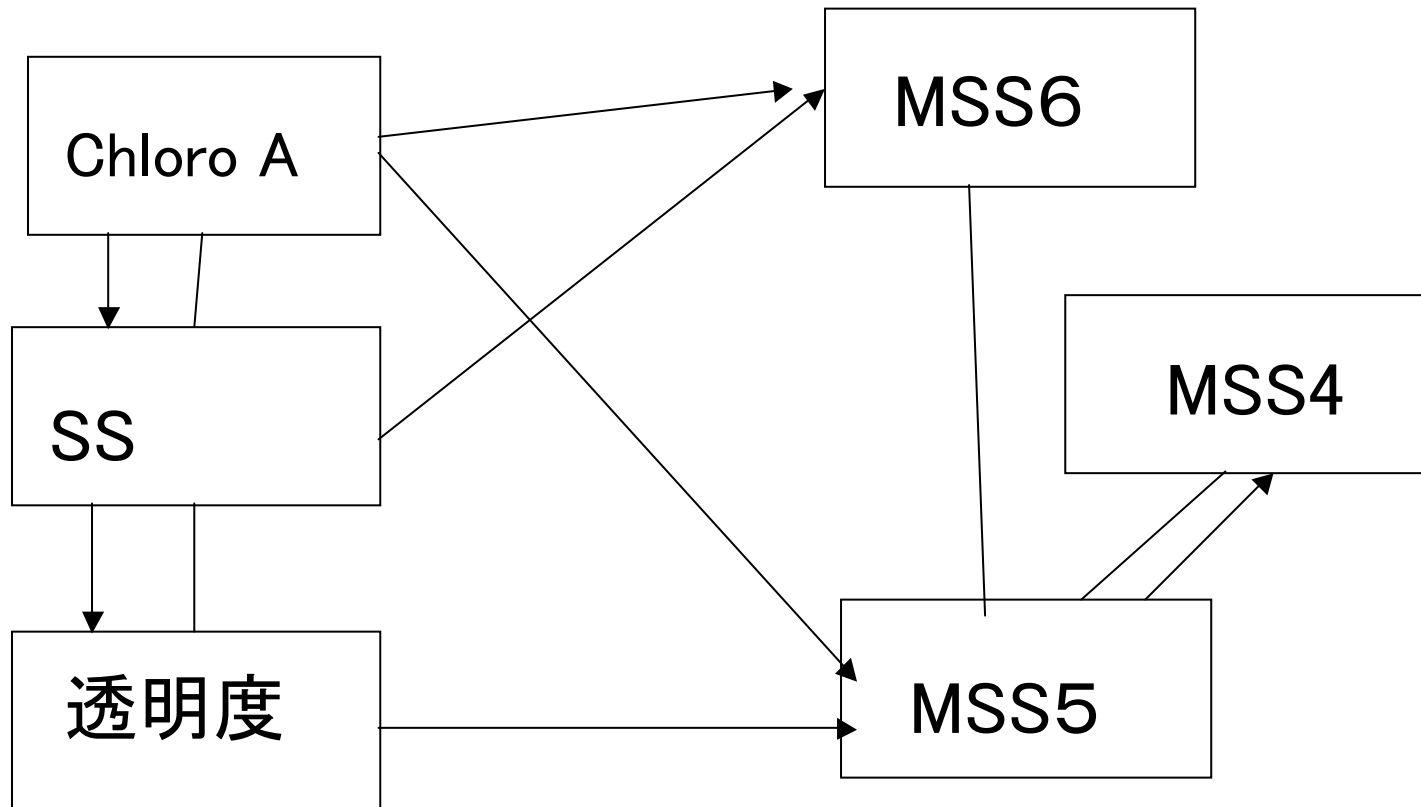


表1 相関係数の2つの推定値
 上三角成分：通常の相関係数、
 下三角成分：適切な単調変換で実現する相関の推定値

	透明度	浮遊懸濁	クロフィル	MSS4	MSS5	MSS6
透明度	1	-0.7719	-0.6782	-0.8646	-0.8536	-0.8095
浮遊懸濁	-0.9604	1	0.7949	0.7713	0.8236	0.7537
クロフィル	-0.8195	0.8405	1	0.6914	0.7212	0.6924
MSS4	-0.8758	0.8812	0.7515	1	0.9819	0.9545
MSS5	-0.8884	0.8905	0.7729	0.9828	1	0.9616
MSS6	-0.8641	0.8594	0.7961	0.9485	0.9696	1

グラフィカルモデリングで予想された因果関係 (探索的因果分析)



因果モデル(標準化)

- $\text{Var}[\text{Chloro}] = \eta_c^2 = 1$
- $\text{SS} = 0.841 \text{ Chloro} + \varepsilon_S, \quad \text{Var}[\varepsilon_S] = \sigma_S^2 = 0.294$
- $\text{Transp} = -0.960 \text{ SS} + \varepsilon_T, \quad \text{Var}[\varepsilon_T] = 0.078$
- $\text{MSS6} = 0.268 \text{ Chloro} - 0.646 \text{ Transp} + \varepsilon_6$
 $\text{Var}[\varepsilon_6] = 0.230$
- $\text{MSS5} =$
 $-0.122 \text{ Chloro} + 0.293 \text{ SS} + 0.818 \text{ MSS6} + \varepsilon_5$
 $\text{Var}[\varepsilon_5] = 0.043$
- $\text{MSS4} = 0.983 \text{ MSS5} + \varepsilon_4, \quad \text{Var}[\varepsilon_4] = 0.034$

共分散構造

$$\Sigma = \begin{bmatrix} \eta_c^2 & 0.841\eta_c^2 & -0.807\eta_c^2 & 0.790\eta_c^2 & 0.770\eta_c^2 \\ 0.841\eta_c^2 & 0.707\eta_c^2 + \sigma_s^2 & -0.679\eta_c^2 - 0.960\sigma_s^2 & 0.664\eta_c^2 + 0.620\sigma_s^2 & 0.648\eta_c^2 + 0.800\sigma_s^2 \\ -0.807\eta_c^2 & -0.679\eta_c^2 - 0.960\sigma_s^2 & 0.0780 + 0.652\eta_c^2 + 0.922\sigma_s^2 & -0.0504 - 0.637\eta_c^2 - 0.595\sigma_s^2 & -0.0412 - 0.622\eta_c^2 - 0.768\sigma_s^2 \\ 0.790\eta_c^2 & 0.664\eta_c^2 + 0.620\sigma_s^2 & -0.0504 - 0.637\eta_c^2 - 0.595\sigma_s^2 & 0.263 + 0.623\eta_c^2 + 0.385\sigma_s^2 & 0.215 + 0.608\eta_c^2 + 0.496\sigma_s^2 \\ 0.770\eta_c^2 & 0.648\eta_c^2 + 0.800\sigma_s^2 & -0.0412 - 0.622\eta_c^2 - 0.768\sigma_s^2 & 0.215 + 0.608\eta_c^2 + 0.496\sigma_s^2 & 0.219 + 0.593\eta_c^2 + 0.640\sigma_s^2 \end{bmatrix}$$

共分散構造を η_c^2 , σ_s^2 の関数で表した.

攪乱母数の関数としての回帰予測

$$Chloro = \frac{\eta_c^2 \left\{ (0.00723 + 0.123\sigma_s^2)MSS6 + (0.0327 - 0.0956\sigma_s^3)MSS5 \right\}}{0.0113 + 0.0309\eta_c^2 + 0.0391\sigma_s^2 + 0.0238\eta_c^2\sigma_s^2}$$

$$SS = \frac{(0.00608\eta_c^2 - 0.0363\sigma_s^2 - 0.0150\eta_c^2\sigma_s^2)MSS6 + (0.0275\eta_c^2 + 0.0769\sigma_s^2 + 0.0413\eta_c^2\sigma_s^2)MSS5}{0.0113 + 0.0309\eta_c^2 + 0.0391\sigma_s^2 + 0.0238\eta_c^2\sigma_s^2}$$

$$Transp = \frac{(-0.0107\eta_c^2 + 0.0230\sigma_s^2 + 0.0144\eta_c^2\sigma_s^2)MSS6 - (0.0214\eta_c^2 - 0.0647\sigma_s^2 - 0.0397\eta_c^2\sigma_s^2)MSS5}{0.0113 + 0.0309\eta_c^2 + 0.0391\sigma_s^2 + 0.0238\eta_c^2\sigma_s^2}$$

クロロフィル信号の分散 $\rightarrow\infty$

$$Chloro = \frac{(0.00723 + 0.123\sigma_s^2)MSS6 + (0.0327 - 0.0956\sigma_s^3)MSS5}{0.0309 + 0.0238\sigma_s^2}$$

$$SS = \frac{(0.00608 - 0.0150\sigma_s^2)MSS6 + (0.0275 + 0.0413\sigma_s^2)MSS5}{0.0309 + 0.0238\sigma_s^2}$$

$$Transp = \frac{(-0.0107 + 0.0144\sigma_s^2)MSS6 - (0.0214 - 0.0397\sigma_s^2)MSS5}{0.0309 + 0.0238\sigma_s^2}$$

水質情報予測式の環境パラメータ依存性
 括弧内左側が MSS6、右側が MSS5 の回帰係数、
 括弧外が予測誤差分散の推定値

η_c	σ_s^2	クロロフィル	SS	透明度
1	0	(0.17, 0.78) 0.268	(0.14, 0.65) 0.189	(-0.30, -0.51) 0.184
<i>1</i>	<i>0.294</i>	<i>(0.72, 0.075)</i> 0.376	<i>(-0.15, 1.03)</i> , 0.220	<i>(-0.030, -0.86)</i> , 0.211
1	∞	(1.96, -1.52), 0.622	(-0.82, 1.88), 0.291	(0.60, -1.66), 0.273
∞	0	(0.23, 1.06), 0.366	(0.20, 0.89), 0.259	(-0.35, -0.69), 0.226
∞	0.294	(1.15, 0.12), 0.602	(0.044, 1.05), 0.265	(-0.17, -0.87), 0.234
∞	∞	(5.19, -4.02), 1.64	(-0.63, 1.74), 0.294	(0.61, -1.67), 0.273

解析結果の考察

- σ_s^2 に大きな値を代入
 - ・ 予測式の形が大きく変化
 - ・ 予測精度も劣化する
- クロロフィルの予測式： σ_s^2 を増大
 - MSS5からMSS6へと重みが移行
 - ・ 予測精度も大きく悪化
 - 一般化が難しいと予想
 - クロロフィルの予測式は、
 - 》クロロフィル増大が、SSに明確な影響を与える環境下では、SSを代用特性として、それと相関性の高いMSS 5を利用して、遠隔計測可能
 - 》クロロフィルとSSの関連性が希薄となる環境下では、むしろクロロフィルと直接パスが引かれるMSS 6を主体に予測式が構成
- このような考察を可能とするのが計測問題における因果推論(順問題同定)の意義

このような柔軟なシミュレーション的 予測が可能なソフトはない

- 任意の予測変数セットから必要な被予測変数の予測式が計算できる共分散構造分析のツールがあれば！！
 - 国産ソフトへの期待
 - 計測技術に答える
 - ◎標本分散ではなく、構造方程式モデリングで推定される共分散構造(Implied Covariance)を用いた最良予測式が計算可能になって欲しい
 - △因子得点推定をBartlett法にする
 - ○環境因子を適宜変えて最良予測式のシミュレーションを可能にする

そして、今、技術の世界でも 仮説的因子が

牧, 沖, 椿, 安岡(2004)

潜在要因の影響を考慮した水質計測の
因果モデリング, 日本リモートセンシング
学会誌, 24巻2号153-162.

- 国立環境研究所開発
高速スペクトルメータに
よる霞ヶ浦水質環境観
測

- 1993年—1996年:29地
点の分光計測

- 440nm,550nm, 675nm,
725nm, 800nm

- クロロフィルA,SSの
水表面同時計測

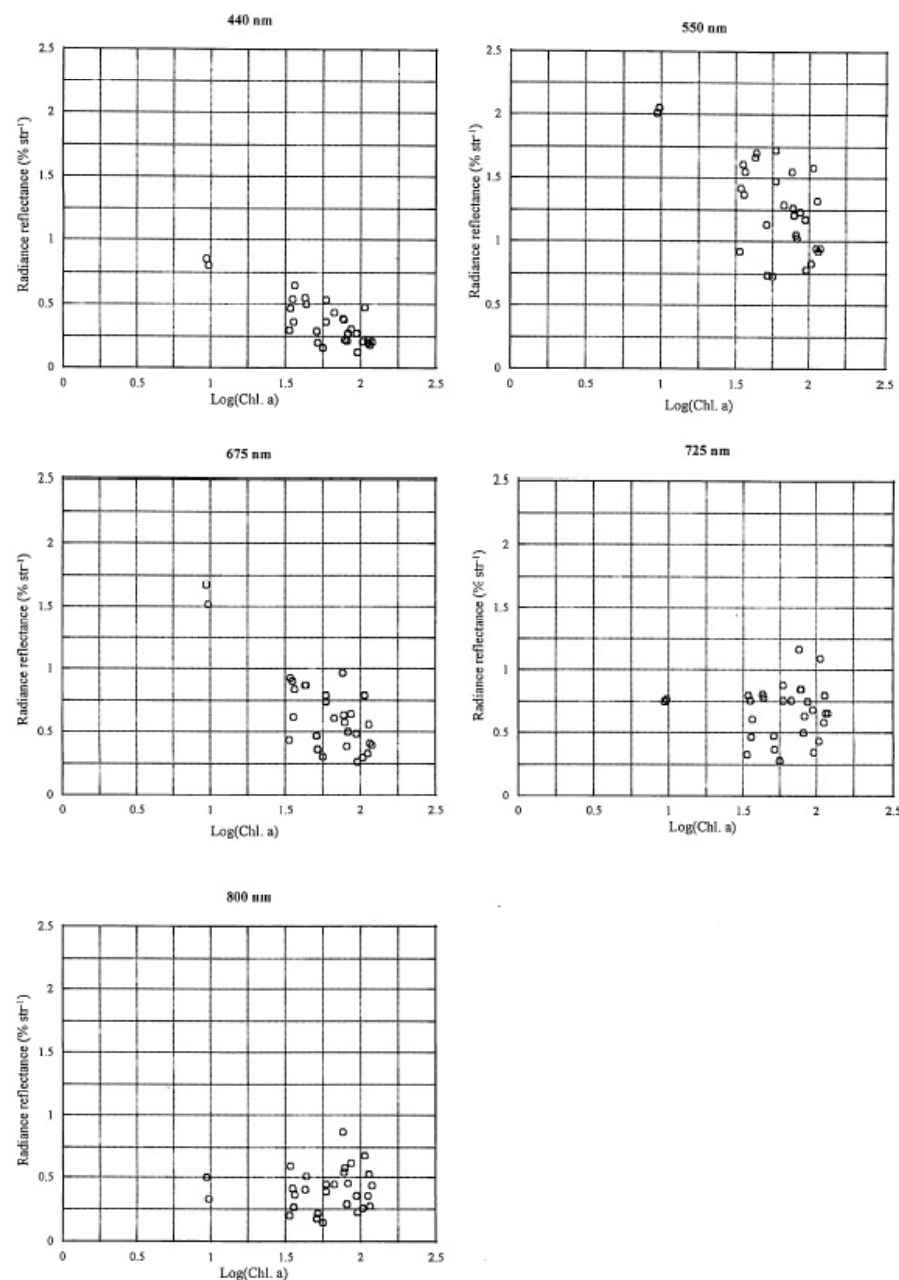


Fig. 3 Relationship between log(Chl. a) and each radiance reflectance.

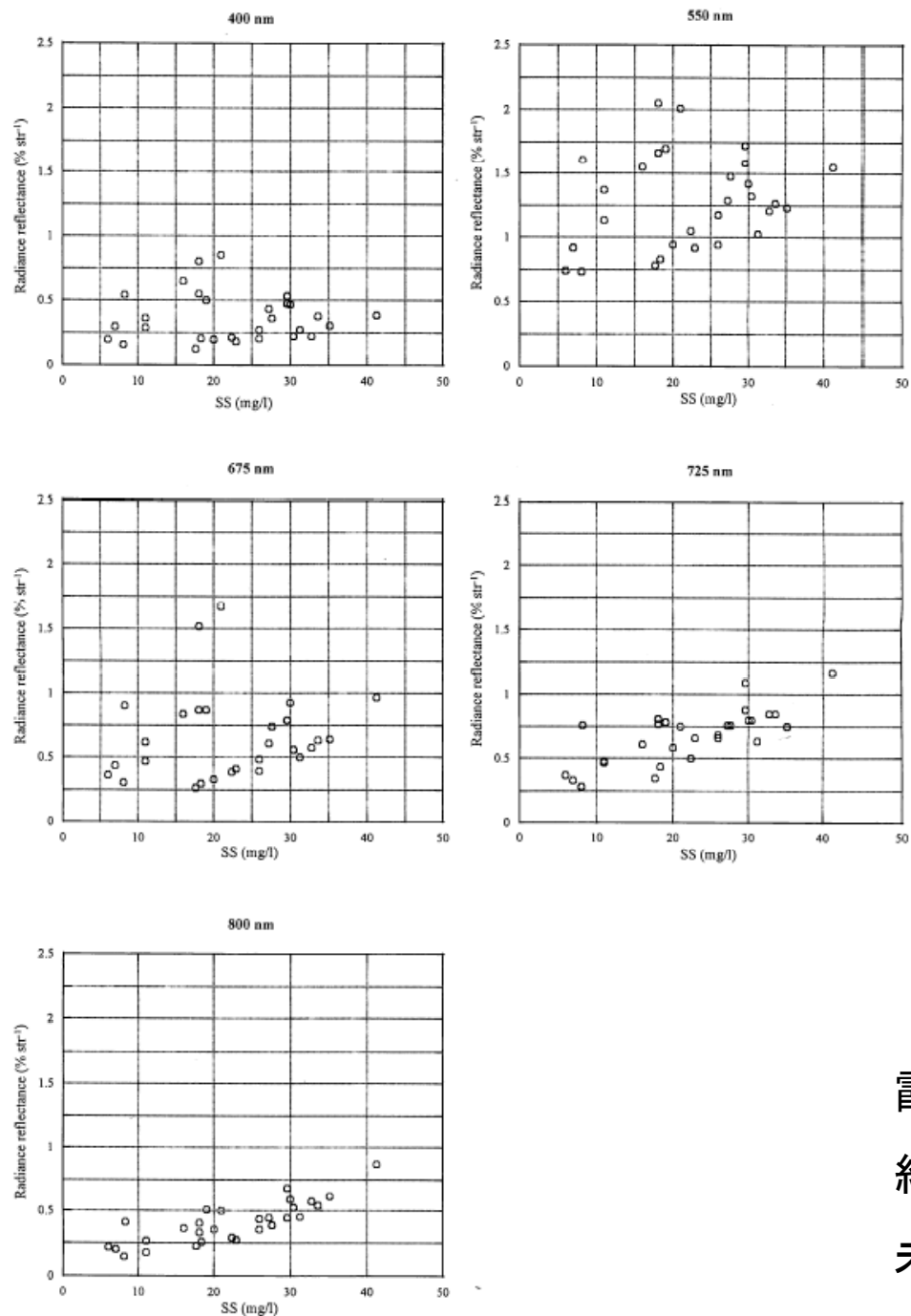


Fig. 4 Relationship between SS concentration and each radiance reflectance.

Table 1 Simple correlation coefficients between spectral data and water quality variables.

	Chl. a	Log(Chl. a)	SS
R _{440nm}	-0.714	-0.803	0.003
R _{550nm}	-0.568	-0.642	0.224
R _{675nm}	-0.655	-0.792	0.155
R _{725nm}	0.066	0.018	0.753
R _{800nm}	0.131	0.101	0.818

Table 2 partial correlation coefficients between spectral data and water quality variables.

	Log(Chl. a)	SS
R _{440nm}	-0.871	0.566
R _{550nm}	-0.808	0.663
R _{675nm}	-0.933	0.813
R _{725nm}	-0.444	0.808
R _{800nm}	-0.401	0.848

霞ヶ浦のSS(浮遊懸濁物質)とは何か？

組成が想像される！

未観測無機物質F1, 未観測有機物質F2

SSが3物質からなるとの仮説モデル 初期仮説と当てはめの改善 (SEMによる独立成分分析)

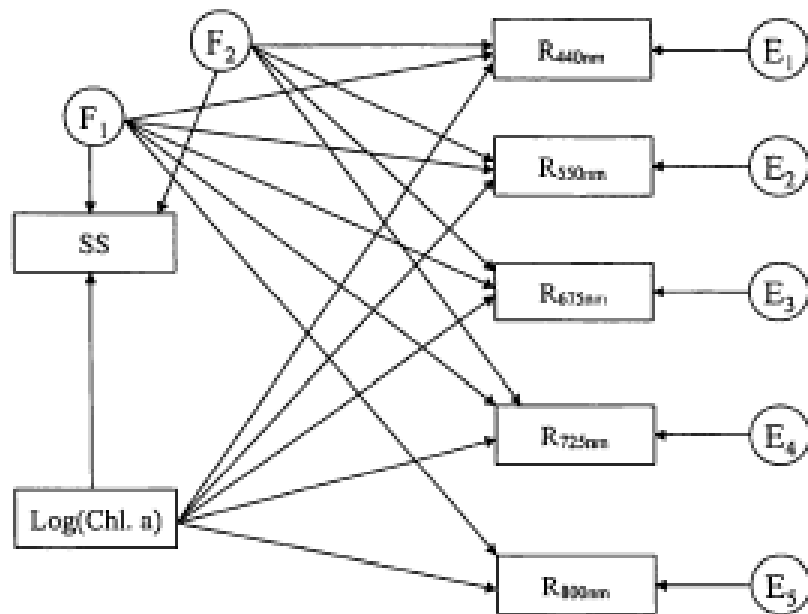
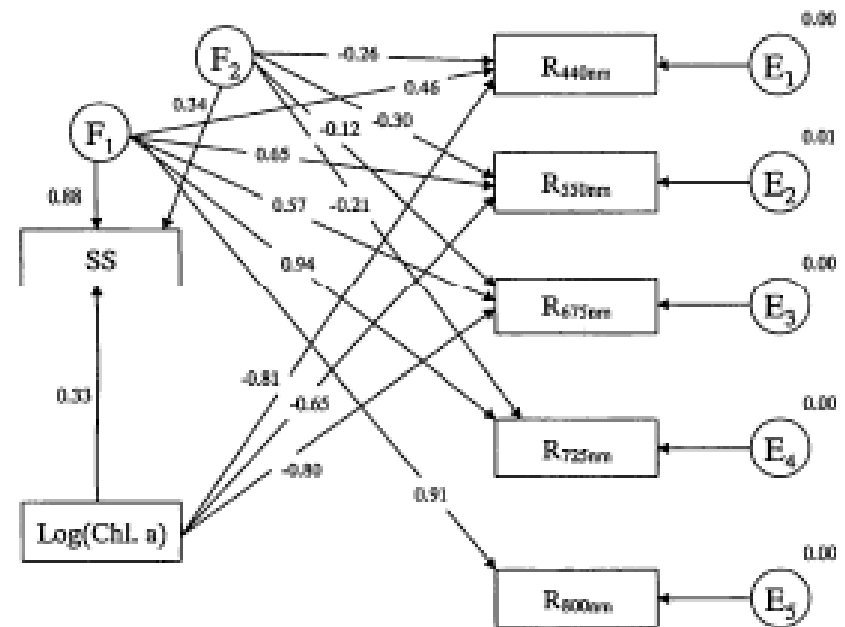


Fig. 6 Path diagram for initial model. E_i ($i = 1 \sim 5$) are error variables for each radiance reflectance. F_i ($i = 1, 2$) are latent variables related to SS characteristic.



Chi-square: 16.962 Degrees of freedom: 7 Probability level: 0.018

Fig. 7 Result of covariance structure analysis to the model which improved Fig. 6.

計測工学からモノづくりへ JUSE多変量解析研究会

- JUSE: 多変量解析研究会での構造方程式モデリング(SEM by AMOS)を紹介
- その後, 多変量解析セミナーのため, TDKの野中氏とQCにおける事例を作成
 - 測定誤差のある回帰分析の事例
 - 厳密な法則が支配しているはずの現象分析の事例
- 因子は真値, 回帰式は法則と考えると共分散構造分析はモノづくりでも必要
 - 心理, マーケティング(日本), BSC(米国)などの文科系的事例とは違った事例を開発し, ノウハウを溜めよう

掲載されている著作物の著作権については、制作した当事者に帰属します。

著作者の許可なく営利・非営利・イントラネットを問わず、本著作物の複製・転用・販売等を禁止します。

所属および役職等は、公開当時のものです。

■公開資料ページ

弊社ウェブページで各種資料をご覧ください <http://www.i-juse.co.jp/statistics/jirei/>

■お問い合わせ先

(株)日科技研 数理事業部 パッケージサポート係 <http://www.i-juse.co.jp/statistics/support/contact.html>