

< 修 士 論 文 >

ガラス流量を制御するための 因果関係の分析

滋 賀 大 学 大 学 院
デ ー タ サ イ エ ン ス 研 究 科
デ ー タ サ イ エ ン ス 専 攻

修了年度：2024年度

学籍番号：6023142

氏 名：馬 渕 美 奈

指導教員：清水 昌平

提出年月日：2025年1月13日

目次

第1章 研究背景.....	1
第2章 目的.....	3
第3章 データと分析手法.....	3
3.1 分析の仮定.....	3
3.2 使用データ.....	5
3.3 親変数の洗い出し.....	7
3.4 因果グラフ.....	8
3.5 モデル作成時に使用する親変数の抽出.....	9
3.6 構造方程式モデルと因果効果.....	10
3.7 感度分析と Robustness value.....	11
第4章 結果.....	13
4.1 因果効果推定.....	13
4.2 感度分析.....	15
第5章 考察.....	17
第6章 結論.....	19
6.1 本研究のまとめ.....	19
6.2 今後の課題.....	19
謝辞.....	20
Appendix.....	21
A1 ハーゲン・ポワズイユの式.....	21
A2 データセットを変更した分析.....	22
A3 仮定を緩めた分析.....	24
A3.1 モデルの線形性とオペレータの調整技量.....	24
A3.2 短期間データを用いた線形モデル分析.....	25
A3.3 長期間データを用いた非線形モデル分析.....	26
A3.4 短期間データを用いた非線形モデル分析.....	28
参考文献.....	29

第1章 研究背景

ガラス製造業においては、24 時間 365 日、高温でガラスを連続的に熔融することにより大量のガラス製品を生産している。これを連続生産という。ガラス製品の連続生産は、大きく分けて熔融、清澄均質化、成形という 3 つの工程から構成されている（図 1.1）。まず、シリカなどの原料を高温で加熱して、液体状のガラス融液にする熔融工程がある。次に、ガラス製品の品質を向上させるための清澄均質化工程が続く。この工程では、ガラス融液中の異物や泡の除去、および均質性向上が図られている。最後に、高温のガラス融液を冷却しながら、目的の形状に成形する成形工程がある。



図 1.1 ガラスの連続生産工程概要図

ガラス製品は最終的な使用用途に応じて様々な形状に成形される。例えば、医薬品などを入れるアンプル管やバイアル瓶に使用される円筒状の管ガラス、スマートフォンやタブレットなどのディスプレイに使用される板状のディスプレイガラスが挙げられる。そのため、成形工程はガラス製品ごとに異なる。しかし、いずれの成形工程においても、高品質なガラス製品を生産するためには成形工程を流れるガラス融液の量（以下、ガラス流量）の安定性が極めて重要な要素となる。ここでは、ディスプレイガラスを例にその理由を説明する。ディスプレイガラスの連続生産工程に関連する品質項目は 9 種類ある。そのうち、ガラス流量の安定性は 6 種類の品質に関わっている。このことから、連続生産工程においてガラス流量の安定性は高品質なガラス製品を生産するための必要条件といえる。

ディスプレイガラスの連続生産方法には、オーバーフロー法やフロート法など複数の製造方法がある（山根 et al. 1999）。本研究では、日本電気硝子株式会社で採用しているオーバーフロー法について取り扱う。図 1.2 にオーバーフロー法によるディスプレイガラスの製造設備の概要図を示す（日本電気硝子株式会社 2018）。

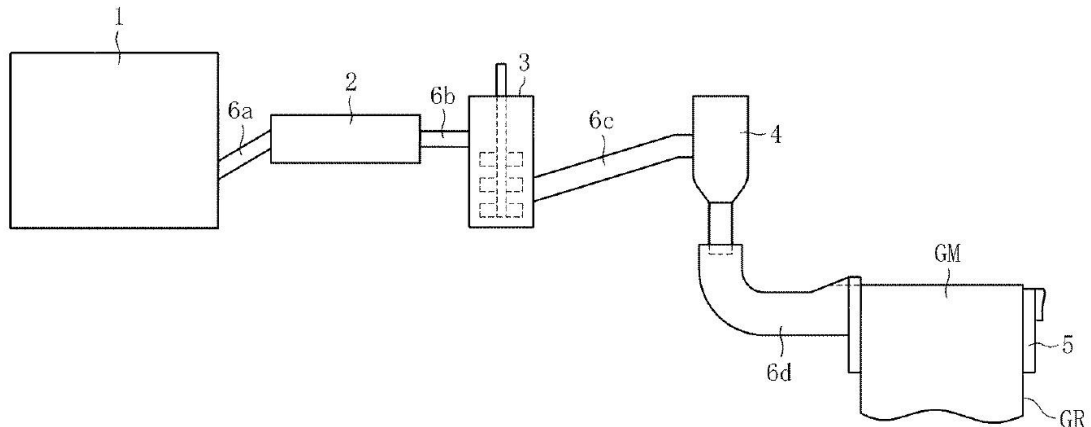


図 1.2 オーバーフロー法によるディスプレイガラス製造設備概要図
 (1：溶解槽、2：清澄容器、3：均質化槽、4：流量調整槽、5：成形槽、
 6a～6d：ガラス供給路、GM：溶融ガラス、GR：ディスプレイガラス)

高品質なディスプレイガラス製品を生産するためには、成形工程、ここでは、成形槽のガラス流量が安定している必要がある。各工程設備では設定温度に維持するための調整機構が常時作動していることに加えて、よりガラス流量を安定させるために、オペレータが成形槽の上流にある流量調整槽の設定温度を調整している。このオペレータによる流量調整槽の設定温度調整は、外気温や上流工程の温度変化を考慮するといった大まかな方針はあるが、具体的にどのタイミングで、どれだけ設定温度を変えるべきかといった、細かい調整方法についてはオペレータの経験や勘に大きく依存している。また、生産中のガラス融液は高温であるため、流量調整槽や成形槽のガラス流量を直接測定することは困難である。そのため、成形後のディスプレイガラス重量を測定することで、間接的にガラス流量を取得することから、流量調整槽の設定温度を調整してから、ガラス流量に影響がでるまでタイムラグがある。オペレータはこのタイムラグも考慮しながら流量調整槽の設定温度を調整する必要がある。これらの理由から、流量調整槽の設定温度の調整技量はオペレータによって差があると言われている。しかし、このオペレータの流量調整槽の設定温度調整がガラス流量にどれだけ影響を与えているかについて、これまで定量的に評価されたことはない。

そこで本研究では、因果推論の枠組みに基づく統計的手法を用いて、流量調整槽の設定温度がガラス流量に与える因果効果を推定した。すなわち、オペレータの介入によるガラス流量への影響を推定し、その評価を試みた。

第2章 目的

本研究では安定して生産が行われている期間（以下、安定生産期間）で発生しているガラス流量の変動について着目した。この安定生産期間において、オペレータによる流量調整槽の設定温度調整がガラス流量へ与えている影響を定量的に評価し、ガラス流量を制御するための知見をオペレータへ提供することを目的とした。

第3章 データと分析手法

3.1 分析の仮定

ここでは本研究で設定した仮定について説明する。説明のために流量調整槽の概要図、および本分析で設定した仮定をまとめたものを図 3.1 に示す。

まず、安定生産期間中にオペレータが調整している流量調整槽の設定温度は2カ所ある。これ以降、その調整箇所を上流から流量調整槽①、流量調整槽②とする。

次に、本分析を行うにあたって設定した仮定について説明する。ガラス流量は「ガラス液面高さ」、および「ガラスの粘度」の2つに大きく支配される。しかし、流量調整槽①のガラス液面高さ、および流量調整槽①、②に流れているガラスの粘度を直接測定することは困難である。そのため、ガラス液面高さは図 1.2 の1溶解槽で測定されているガラス液面高さで代用することとした。一方で、ガラスの粘度は流量調整槽自体の温度で代用することとした。工程設備内を流れるガラスの粘度はガラスの温度と指数関数の関係にあることがわかっており（松岡 2017）、ガラスの温度と流量調整槽自体の温度は同じと仮定すると、ガラスの粘度と流量調整槽自体の温度も指数関数の関係にあるといえると考えたためである。

本研究では安定生産期間のデータを用いるため、ガラス流量の変動幅は非常に小さく、指数関数の中の微小区間のデータを使用していると見なすことができる（図 3.2）。そのため、ガラスの粘度と流量調整槽自体の温度の関係性を線形で近似できるとして、以降の分析を行った。

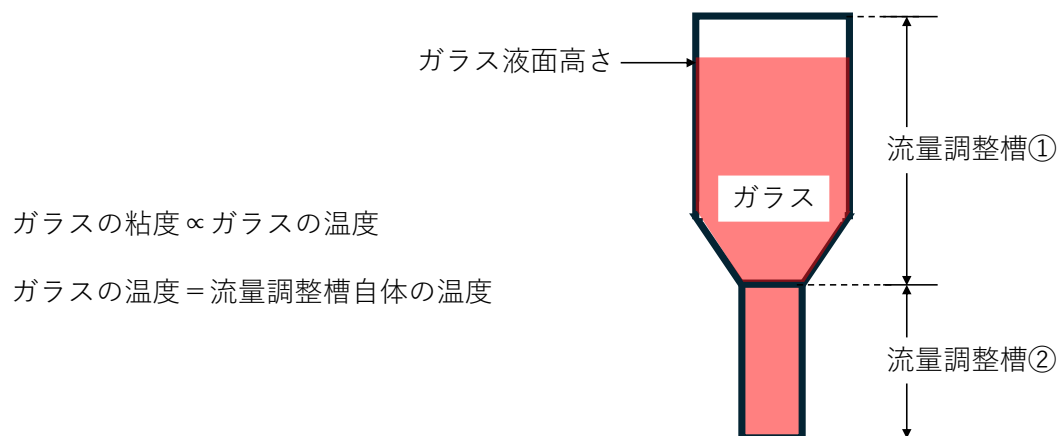


図 3.1 流量調整槽概要図と本分析で設定した仮定

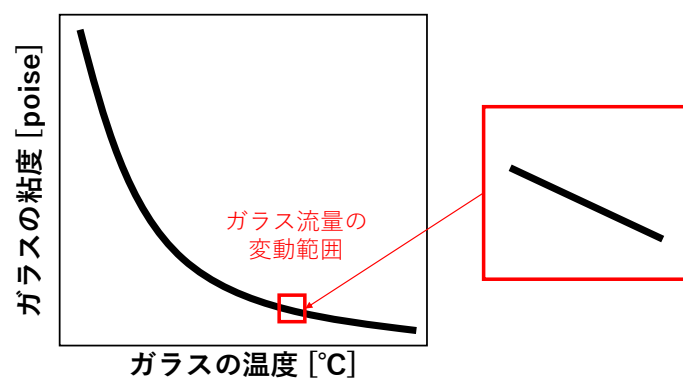


図 3.2 安定生産期間中のガラス流量変動範囲における線形性の概要図

3.2 使用データ

各工程設備では 1 分間の取得頻度でデータが自動的に記録されている。オペレータへのヒアリングを行い、本研究で安定生産期間として抽出する判断基準を以下にまとめた。

- A) ガラス流量が 2 分以上欠損していないこと。
- B) ガラス流量とその 1 階差分を時系列プロットで可視化したとき、明らかな外れ値がないこと。
- C) ガラス組成変更のような大幅な生産条件の変更がないこと。
- D) ガラスの均質性に基づく品質項目の値が大きく変化していないこと。
- E) ガラス液面高さを補正する作業が実施されていないこと。

項目 A については、1 分間のみの欠損値があった場合は前後の値から線形補完して使用した。その他のデータについても、欠損値があった場合は線形補完を行って使用することとした。項目 B を確認することで、ガラス流量の誤差の期待値が 0、かつ分散が一定という仮定を満たしていると考えられる。項目 C～E は、期間内で生産条件が一定である仮定を満たすために設定した。

これらの項目を満たす期間を安定生産期間として抽出した結果、4 つのデータセットを作成できた。その中から最も観測数が多い期間（観測数=18,572）を用いて今後の分析を行うこととした。

図 3.3 に分析で使用した期間のガラス流量、および流量調整槽①、②設定温度のグラフを示す。データの機密性から、縦軸の具体的な数値については表記してない。図 3.4 には同じ変数の 1 階差分のグラフを示す。

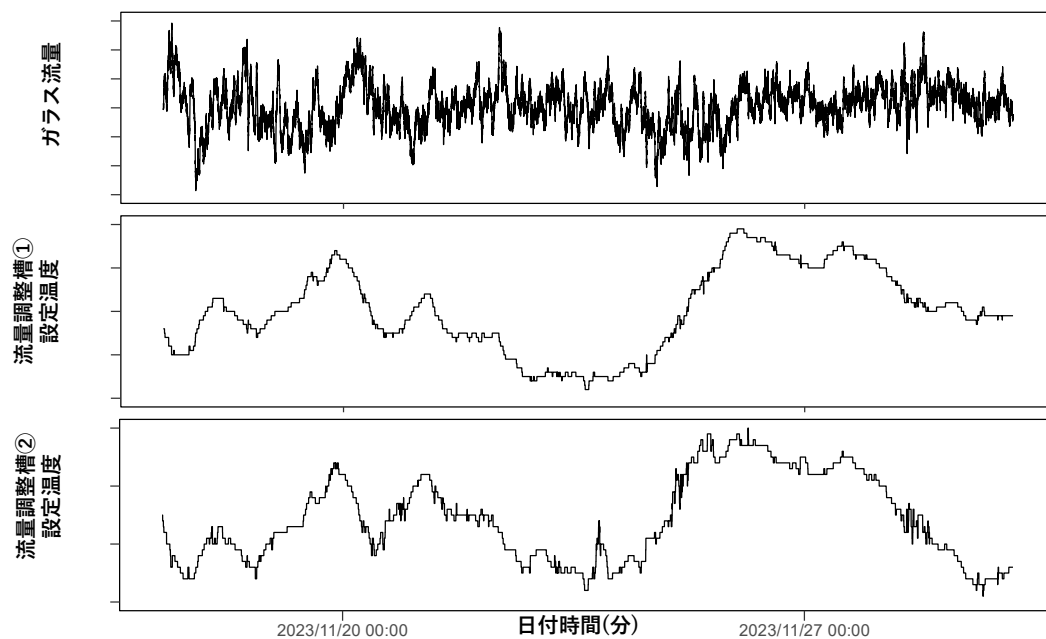


図 3.3 安定生産期間中の各変数の時系列プロット
(上：ガラス流量、中：流量調整槽①設定温度、下：流量調整槽②設定温度)

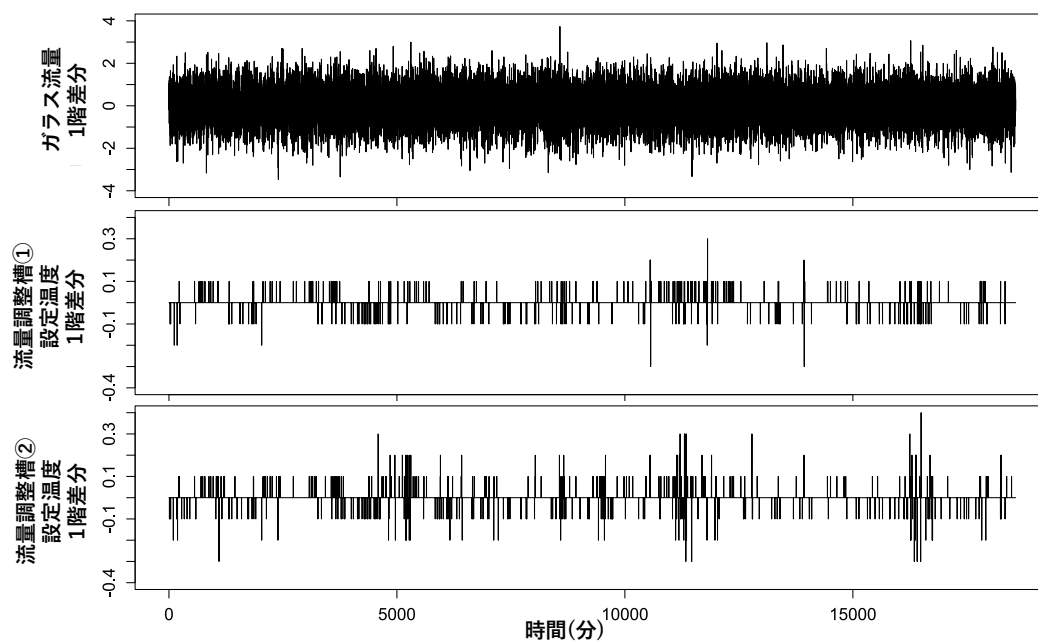


図 3.4 安定生産期間中の各変数の1階差分時系列プロット
(上：ガラス流量、中：流量調整槽①設定温度、下：流量調整槽②設定温度)

3.3 親変数の洗い出し

ガラス流量と流量調整槽①、②の設定温度に関する因果グラフを作成するため、オペレータにヒアリングを行い、流量調整槽の設定温度を調整する際に参考になっている変数を調査した。得られた変数の一覧を表 3.1 に示す。

表 3.1 ヒアリングにより選定した流量調整槽①、②の設定温度の親変数一覧

設備 A_2_電流値_2分傾き	設備 B_北温度_2分傾き
設備 A_2_電流値_10分傾き	設備 B_北温度_10分傾き
設備 A_3_電流値	設備 B_下温度_2分傾き
設備 A_3_電流値_2分傾き	設備 B_下温度_10分傾き
設備 A_3_電流値_10分傾き	設備 B_上温度_2分傾き
設備 A_4_電流値_2分傾き	設備 B_上温度_10分傾き
設備 A_4_電流値_10分傾き	設備 B_下温度_2分傾き
流量調整槽①_電流値	設備 B_下温度_10分傾き
流量調整槽①_電流値_2分傾き	室外気温
流量調整槽①_電流値_10分傾き	室外気温_10分傾き
流量調整槽②_電流値	室外湿度
流量調整槽②_電流値_2分傾き	室外湿度_10分傾き
流量調整槽②_電流値_10分傾き	現在温度と設定温度の差（流量調整槽①／流量調整槽②）
設備 A_2_先温度_2分傾き	偏肉_傾き
設備 A_2_先温度_10分傾き	ガラス液面高さ
設備 A_2_元温度_2分傾き	ガラス流量
設備 A_4_元温度_10分傾き	ガラス流量_2分傾き
設備 A_4_先温度_2分傾き	ガラス流量_10分傾き
設備 A_4_先温度_10分傾き	
流量調整槽①_下東温度_2分傾き	
流量調整槽①_下東温度_10分傾き	
流量調整槽②_南温度_2分傾き	
流量調整槽②_南温度_10分傾き	
設備 B_南温度_2分傾き	
設備 B_南温度_10分傾き	

設備温度、および電流値の 2 分傾きの変数については、2 分前のデータとの差をとって、2 で割って算出した。また、設備温度、および電流値の 10 分傾きの変数については、オペレータは微小な変動を無視しておおよその傾きを見ているため、データの前後 1 点ずつを含む 3 点移動平均値を算出した後、10 分前とのデータとの差を取り、10 で割って使用することとした。偏肉の傾きの変数はデータの前後 2 点ずつを含む 5 点移動平均を算出した後、差分をとって使用することとした。

3.4 因果グラフ

分析データは時系列データのため、時間遅れの効果も併せて検討した。ガラス流量を取得する際に発生するタイムラグを考慮し、ガラス流量は Lag 分だけずらしたデータを使用した。本研究で使った流量調整槽①設定温度(t)からガラス流量($t + \text{Lag}$)への因果グラフの概要図を図 3.5 に示す。すべての親変数を表示することは難しいため、親変数は一部のみ描画した。流量調整槽②についても同様の因果グラフを使用した。

この因果グラフから、本分析ではガラス流量の自己回帰を含めた $\text{ARX}(p)$ モデル (Auto-Regressive eXogenous model) (Ljung 1999; 足立 1993) を作成し、流量調整槽の設定温度からガラス流量への因果効果を推定することとした。

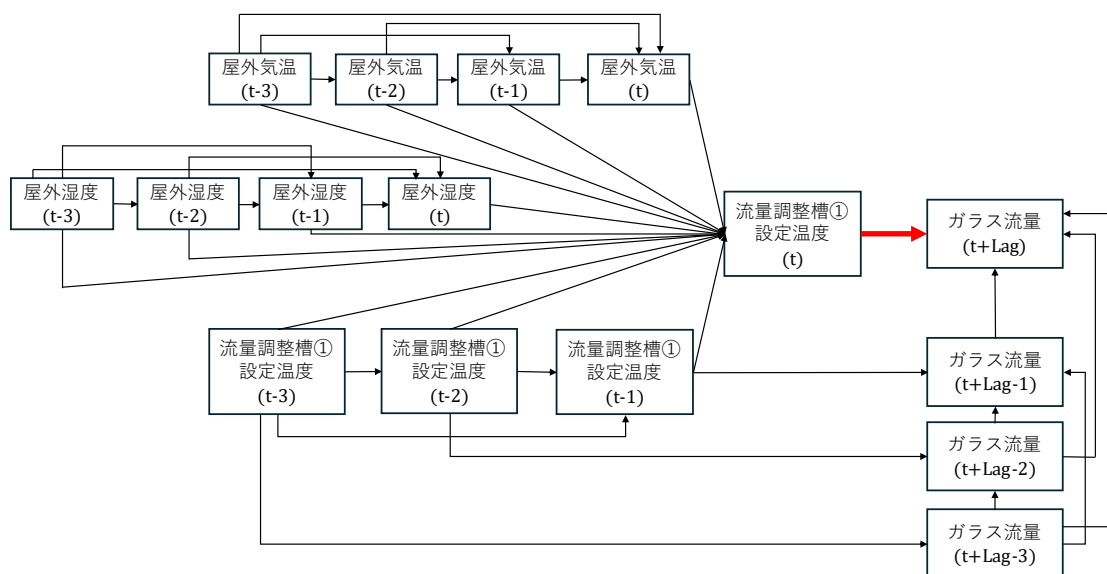


図 3.5 作成した因果グラフ ($p=3$ の時、親変数は屋外気温と屋外湿度のみ描画)

3.5 モデル作成時に使用する親変数の抽出

オペレータは各設備の温度や電流値を参考にしながら、流量調整槽の設定温度を調整している。これらの変数の中には、1つの設備に対して異なる位置に設置されている測定点から得られたものがある。例えば、流量調整槽①には、異なる位置に温度を測定するための熱電対が5点設置されている。これは、ある測定点が劣化や故障などによって正常なデータを記録できない状態になった時、他の測定点で代用するためである。オペレータは同一設備に設置された複数の測定点のデータを見て各測定点のデータに異常がないか確認しながら、流量調整槽①、②の設定温度を調整している。そのため、モデルを作成する際、1つの設備の中でどの測定点をモデルに使用するのが最も適切か判断することは困難であった。しかし、ほぼ同一の測定点と見なせる変数、つまり、多重共線性のある変数同士をモデルに入れると、標準誤差が大きくなって推定が正確に行われなくなる恐れがある。そこで本分析では $VIF > 10$ の変数があった場合に一方の変数を除外し、多重共線性を回避することとした。

また、自己相関が高い変数があったため、一部の変数は測定データの1階差分を使用することとした。

これらをふまえて、モデル作成時に使用した変数名の一覧を表3.2に示す。

表 3.2 推定に使用した流量調整槽①、②の設定温度の親変数一覧

設備A_2_電流値_2分傾き	設備B_元北温度_2分傾き
設備A_3_電流値_1階差分	設備B_元北温度_10分傾き
設備A_3_電流値_2分傾き	設備B_中下温度_2分傾き
設備A_4_電流値_2分傾き	設備B_中下温度_10分傾き
流量調整槽①_電流値_1階差分	設備B_先上温度_2分傾き
流量調整槽①_電流値_2分傾き	設備B_先上温度_10分傾き
流量調整槽②_電流値_1階差分	設備B_先下温度_2分傾き
流量調整槽②_電流値_2分傾き	設備B_先下温度_10分傾き
設備A_2_先温度_2分傾き	室外気温_1階差分
設備A_2_先温度_10分傾き	室外湿度_1階差分
設備A_4_元温度_2分傾き	現在温度と設定温度の差(流量調整槽①/流量調整槽②)
設備A_4_元温度_10分傾き	偏肉_傾き
設備A_4_先温度_2分傾き	ガラス液面高さ_1階差分
設備A_4_先温度_10分傾き	ガラス流量_1階差分
流量調整槽①_下東温度_2分傾き	ガラス流量_2分傾き
流量調整槽①_下東温度_10分傾き	
流量調整槽②_南温度_2分傾き	
流量調整槽②_南温度_10分傾き	
設備B_元南温度_2分傾き	
設備B_元南温度_10分傾き	

3.6 構造方程式モデルと因果効果

Sewall Wright は因果効果を定量的に評価するための方法論としてパス解析 (path analysis) (Wright 1921) を提案した。パス解析は遺伝学の分野で開発され、その後構造方程式モデリング (Structural Equation Modeling : SEM) (Wold 1954) として、計量心理学や計量経済学の分野で発展した。SEM はデータ生成過程と呼ばれる変数の値が決定される順序を数学的に表現している。例として、 x と y の 2 変数のデータ生成過程を考える。この場合のデータ生成過程を構造方程式 (3.1)、(3.2) を用いて表現すると、以下のように表される。

$$x = e_x \quad (3.1)$$

$$y = f_y(x, e_y) \quad (3.2)$$

構造方程式は両辺が等しいだけでなく、左辺が右辺で決定される因果関係も表している。つまり、式 (3.2) の y は、 x と e_y の値から関数 f_y を通して決まることが表現されている。 y は内生変数 (endogenous variable)、誤差変数 e_x 、および e_y は外生変数 (exogenous variable) と呼ばれる。外生変数はそれぞれの値を決定するために寄与しうる未観測変数も含まれているため、構造方程式では右辺に登場する。

因果効果は主に 3 つに分類される。ある変数へ直接的に与える因果効果を「直接因果効果」、他の変数を経由して間接的に与える因果効果を「間接因果効果」と呼ぶ。そして、この 2 つを合わせた因果効果を「総合因果効果」と呼ぶ (清水 2006)。本研究では、オペレータが調整している流量調整槽の設定温度からガラス流量への直接的な因果効果を評価するため、直接因果効果を本研究における「因果効果」とする。

構造方程式では変数間の関係に線形性が仮定された場合、その構造方程式は線形構造方程式モデル (黒木 2017) と呼ばれる。このモデルにおいて変数間の因果関係が正しく記述されているとき、データに基づいて因果効果を推定できる場合がある。この時の因果効果は各変数 x_i ($i = 1, 2, \dots, p$) に関する線形構造方程式に基づき推定できる。

$$x_i = \sum_{k(j) < k(i)} b_{ij} x_j + e_i \quad (3.3)$$

右辺の係数 b_{ij} は、左辺の変数 x_i を目的変数として扱い、左辺の直接的な原因とされる親変数の候補、すなわち因果的順序が $k(j) < k(i)$ となるような変数を用いて回帰分析をした時の x_j から x_i への偏回帰係数で推定できる。本研究では変数同士に線形性を仮定したため、線形重回帰分析において推定した偏回帰係数を因果効果とした。データはすべて連続値として扱ったため、最小二乗法を用いて偏回帰変数を推定した。

3.7 感度分析と Robustness value

観察データを用い因果推論的手法によって因果効果を推定する分析では、基本的に未観測共通原因がないという仮定が置かれている。使用したデータや分析手法に適した感度分析により、未観測共通原因が研究デザインに影響を及ぼしていないことを示すことで、この仮定を満たしているか確認できる。一般的に感度分析は分析時に用いた仮定のうち、1つを緩和した場合に結果がどの程度変わるかを評価する手法の総称だが、本研究ではこれ以降、感度分析は未観測共通原因に対する感度分析を指すものとする。

感度分析では、推定結果に対して未観測共通原因がどのくらい影響力を持つのか、また、どの程度の強さの未観測共通原因があれば分析結果に影響を及ぼすかといったベンチマークを示す必要がある。

測定されていない未観測共通原因について何も情報がない場合に、部分決定係数 (partial R^2) を調べることで、未観測共通原因が因果効果の推定結果をどのように変えうるか確認できる。目的変数を Y 、目的変数 Y に対する因果効果を推定したい変数を X (以下、処置 X とする)、共変量を Z 、未観測共通原因を U とすると、部分的決定係数は式 (3.4)、(3.5) で算出できる。この時、 $R_{Y \sim X+Z}^2$ は目的変数 Y に対して処置 X と共変量 Z を用いて線形重回帰分析をしたときの決定係数である。

$$R_{Y \sim U|X,Z}^2 = \frac{R_{Y \sim U+X+Z}^2 - R_{Y \sim X+Z}^2}{1 - R_{Y \sim X+Z}^2} \quad (3.4)$$

$$R_{X \sim U|Z}^2 = \frac{R_{X \sim Z+U}^2 - R_{X \sim Z}^2}{1 - R_{X \sim Z}^2} \quad (3.5)$$

$R_{Y \sim U|X,Z}^2 = 1$ の場合、すなわち、単一の未観測共通原因 U が、目的変数 Y に対して処置 X と共変量 Z を用いて線形重回帰分析をした時の残差分散のうち、100%を説明できる強さを持つ場合、推定された処置 X の偏回帰係数は0になる。この時、 $R_{X \sim U|Z}^2 = R_{Y \sim X|Z}^2$ であることも意味しており、極端に強い未観測共通原因 U が、目的変数 Y に対して処置 X と共変量 Z を用いて線形重回帰分析をした時の残差分散のうちの100%を説明できるとしても、推定された処置 X の偏回帰係数を0にするためには、その未観測共通原因 U が処置 X とどの程度強く関連している必要があるか、ということも示している。

Cinelli と Hazlett は、この部分決定係数を利用して、連続値を用いた場合の因果効果の推定結果の信頼性を評価できる Robustness value (3.6) という指標を提案した (Cinelli & Hazlett 2020)。

$$RV = \frac{1}{2} \left(\sqrt{f^4 + 4f^2} - f^2 \right) \quad (3.6)$$

f は目的変数Yと処置Xの部分コーヘンの f (partial Cohen's f) である。また、 f は処置Xの偏回帰係数の t 値を自由度の平方根 (\sqrt{df}) で割ることでも算出できる。

RV は、 $R_{Y \sim U|X,Z}^2 = R_{X \sim U|Z}^2 = RV$ であることから、推定された処置Xの偏回帰係数を 0 にするためには、未観測共通原因Uは目的変数Yと処置Xの両方に対してどれだけ関連している必要があるかを示している。

また、有意水準 α 、および自由度 $df - 1$ における t 値の閾値を $t_{\alpha, df-1}^*$ と表した時、 $f_{df-1}^* = |t_{\alpha, df-1}^*|/\sqrt{df-1}$ として、(3.7) 式で算出される f_α を用いて RV_α (3.8) を算出できる。この RV_α は、 α の有意水準で統計的に有意となる境界を示している。この境界を超える強さを持つ未観測共通原因があった場合、 α の有意水準で因果効果の推定結果が 0 でないとはいえない。

$$f_\alpha = f - f_{df-1}^* \quad (3.7)$$

$$RV_\alpha = \frac{1}{2}(\sqrt{f_\alpha^4 - 4f_\alpha^2} - f_\alpha^2) \quad (3.8)$$

本研究では、目的変数Yはガラス流量、処置Xは流量調整槽①、②の設定温度、共変量Zはオペレータが調整時に考慮している変数である。

第4章 結果

4.1 因果効果推定

本分析ではガラス流量を測定する際に発生するタイムラグ（Lag）を考慮し、流量調整槽の設定温度調整から7~16分後の各10時点でガラス流量への因果効果を推定することとした。多重検定の補正には、ボンフェローニ補正（Bonferroni 1936）を用いた。本分析では5%を有意水準として設定し、検定回数の10で割った0.005を有意水準として用いて統計的に有意か判断した。

ARX(p)モデルの時間遅れ次数 p は BIC を用いて選定した。 $p=1\sim5$ の範囲でそれぞれのLagにおいて最適な時間遅れ次数 p を使用した。

最小二乗法による因果効果の推定後、残差(t)と残差(t-p)の散布図とその相関係数により残差に系列相関が残っていないか確認した。相関係数はいずれの場合でも絶対値が0.1以下と十分低く、残差に系列相関は残っていないと判断した。例として、Lag=10の場合を示す。流量調整槽①、②の設定温度はどちらも $p=3$ が最適な次数として選定されたため、ARX(3)モデルを用いて推定を行った。残差(t)と残差(t-3)の相関係数を確認したところ、どちらも-0.02と十分小さかった。その散布図を図4.1に示す。

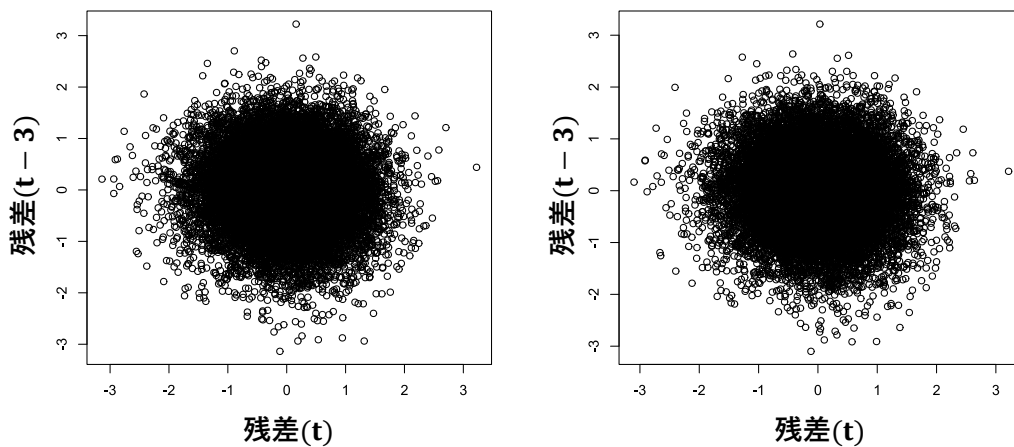


図 4.1 Lag=10 の時の残差(t)と残差(t-3)の散布図

(左：流量調整槽①設定温度の場合、右：流量調整槽②設定温度の場合)

流量調整槽①設定温度からガラス流量への因果効果を推定した結果を表4.1に示す。

分析の結果、流量調整槽①設定温度の調整から8~16分後の9時点で有意な因果効果が確認された。その中で最も因果効果が大きかったのは、調整から13分後だった。その13分後からLagが離れるほど因果効果が小さくなっていることから、流量調整槽①設定温度の調整は時間的な幅を持ってガラス流量に影響を与えていることが確認できた。

表 4.1 流量調整槽①設定温度からガラス流量への因果効果推定結果まとめ

Lag	p	β	S. E.	95%信頼区間		調整済み R^2
				Lower	Upper	
7	2	0.30	0.38	-0.44	1.04	0.28
8	4	1.43	0.38 *	0.69	2.16	0.30
9	4	1.59	0.38 *	0.85	2.32	0.31
10	3	1.77	0.38 *	1.03	2.51	0.29
11	2	1.58	0.38 *	0.83	2.32	0.27
12	1	2.05	0.38 *	1.30	2.80	0.26
13	4	2.44	0.38 *	1.70	3.19	0.29
14	4	2.25	0.38 *	1.51	3.00	0.29
15	3	2.11	0.38 *	1.36	2.86	0.27
16	2	1.46	0.38 *	0.71	2.21	0.27

Note. * <0.005

次に、流量調整槽②設定温度からガラス流量への因果効果を推定した結果を表 4.2 に示す。

表 4.2 流量調整槽②設定温度からガラス流量への因果効果推定結果まとめ

Lag	p	β	S. E.	95%信頼区間		調整済み R^2
				Lower	Upper	
7	2	1.65	0.29 *	1.08	2.21	0.28
8	4	2.03	0.29 *	1.47	2.60	0.30
9	4	2.15	0.29 *	1.59	2.72	0.30
10	3	1.79	0.29 *	1.22	2.36	0.28
11	2	1.29	0.29 *	0.72	1.86	0.27
12	1	1.44	0.29 *	0.87	2.01	0.26
13	4	1.57	0.29 *	1.00	2.14	0.29
14	4	1.35	0.29 *	0.78	1.92	0.29
15	3	1.10	0.29 *	0.53	1.67	0.27
16	2	0.88	0.29 *	0.31	1.45	0.27

Note. * <0.005

この結果から、流量調整槽②設定温度では調整から 7~16 分後のすべての時点で有意な因果効果が確認された。最も因果効果が大きかったのは調整から 9 分後だった。流量調整槽①の推定結果と同様に、9 分後を頂点に時間が離れるにつれて因果効果が小さくなっていることから、時間的な幅を持ってガラス流量に影響を与えていることが確認できた。

最も因果効果が大きかった Lag は、流量調整槽①設定温度は 13 分後、流量調整槽②設定

温度は 9 分後だった。この Lag の差は、設備構造の差と考えられる。流量調整槽②のほうがより下流側、すなわち、よりガラス流量の測定点に近いため、最大の因果効果の Lag が小さくなったと考えられる。

流量調整槽①、②の設定温度どちらの場合でも、Lag によって最適な次数 p は異なった。ガラス流量は目標値を上下しており、一部期間では周期性があるように見られることから、その周期性与合う Lag では次数 p が小さく、その周期とずれる Lag では、次数 p が大きくなる傾向だったのではないかと考えている。

最後に、流量調整槽①、②の設定温度からガラス流量への効果は必ず正であることから、この推定結果は最低限の信頼性が保証されていると考えている。

4.2 感度分析

これまでの分析では未観測共通原因 U がないという仮定が置かれているが、本研究で得られた推定結果が有効であるかを確認するため、感度分析を行った。流量調整槽①設定温度の因果効果について、最も因果効果が大きかった $\text{Lag} = 13$ の場合における感度分析の結果を表 4.3 に示す。

表 4.3 流量調整槽①設定温度の感度分析結果まとめ

	β	S. E.	t-value	$R_{Y \sim X Z}^2$	RV	$RV_{\alpha=0.005}$
流量調整槽① 設定温度	2.44	0.38	6.46	0.2%	4.6%	2.6%

df = 18379, Bound(U が偏肉傾き(t-4)と同じ強さの時) : $R_{Y \sim U|X,Z}^2 = 3.8\%$, $R_{X \sim U|Z}^2 < 0.1\%$

まず流量調整槽①設定温度(処置 X) では $RV = 4.6\%$ と算出された。これは、目的変数 Y であるガラス流量を流量調整槽①設定温度と共変量 Z で線形重回帰分析した時の残差分散のうち、 4.6% を未観測共通原因 U が説明する場合、推定した流量調整槽①設定温度の偏回帰係数が 0 になることを示す。また、 $RV_{\alpha=0.005} = 2.6\%$ が $R_{Y \sim U|X,Z}^2$ より大きければ、もし未観測共通原因 U が存在しても、有意水準 α において推定結果は統計的に有意であると言える。そして、もしガラス流量に対して流量調整槽①設定温度と共変量 Z で回帰した時の残差分散のうち、 100% を説明できる未観測共通原因 U があったとしても、推定された流量調整槽①設定温度の偏回帰係数を 0 にするためには、流量調整槽①設定温度に対して未観測共通原因 U と共変量 Z を用いて回帰した時の残差分散のうち、少なくとも $R_{Y \sim X|Z}^2 = 0.2\%$ を説明する必要があることを示している。

未観測共通原因 U については情報がないため、共変量 Z の中で最も効果が大きかった偏肉傾き(t-4)で代用することとした。もし未観測共通原因 U が偏肉傾き(t-4)と同じだけの強さを持っていた場合、 $R_{Y \sim U|X,Z}^2 = 3.8\%$ と算出された。この値は $RV = 4.6\%$ より小さいため、

もしそのような未観測共通原因 U を見落としていたとしても、推定された流量調整槽①設定温度の偏回帰係数が 0 になることはない。また、 $R_{X \sim U|Z}^2$ は $R_{Y \sim U|X,Z}^2 = 0.2\%$ より小さいため、ガラス流量を流量調整槽①設定温度とその共変量 Z で回帰した時の残差分散を100%説明できるような強力な未観測共通原因 U があったとしても、推定結果が 0 になることはないといえる。その一方で、 $R_{Y \sim U|X,Z}^2 = 3.8\%$ は $RV_{\alpha=0.005} = 2.6\%$ より大きいので、もし偏肉傾き($t-4$)と同じ強さを持つ未観測共通原因 U が存在した場合、本研究で設定した0.5%有意水準では推定結果は統計的に有意でない、すなわち、推定された偏回帰係数が 0 でないとはいえない。

次に、流量調整槽②設定温度について、最も因果効果が大きかったLag = 9の場合の感度分析の結果を表 4.4 に示す。

表 4.4 流量調整槽②設定温度の感度分析結果まとめ

	β	S. E.	t-value	$R_{Y \sim X Z}^2$	RV	$RV_{\alpha=0.005}$
流量調整槽② 設定温度	2.15	0.29	7.49	0.3%	5.4%	3.4%

df = 18416, Bound(U が偏肉傾き($t-3$)と同じ強さの時) : $R_{Y \sim U|X,Z}^2 = 2.7\%$, $R_{X \sim U|Z}^2 < 0.1\%$

流量調整槽②設定温度においても同様に考えることができる。流量調整槽②設定温度においては、共変量 Z の中で最も効果が高い変数は偏肉傾き($t-3$)であった。この変数と同じ強さの未観測共通原因 U が存在した場合、 $R_{Y \sim U|X,Z}^2 = 2.7\%$ と算出された。これは、 $RV = 5.4\%$ 、および $RV_{\alpha=0.005} = 3.4\%$ のどちらよりも低かった。このことから、偏肉傾き($t-3$)と同じ強さの未観測共通原因 U を見落としていたとしても、推定結果が完全に 0 にならず、また、0.5%有意水準においても統計的に有意であるといえる。また、 $R_{X \sim U|Z}^2$ は $R_{Y \sim U|X,Z}^2 = 0.3\%$ より小さく、もしガラス流量に対して流量調整槽②とその共変量 Z で線形重回帰分析した時の残差分散のすべてを説明できるような強力な未観測共通原因 U があったとしても、推定結果が 0 になることはないことがわかった。

感度分析の結果から、共変量 Z の中で最も効果が高い変数と同じ強さを持つ未観測共通原因 U を見落としていたとしても、本研究で得られた推定結果は 0 にならないことが確認できた。また、そのような強力な未観測共通原因 U を熟練のオペレータが見落としている可能性は非常に低いと考えられることから、本研究の推定結果が完全に 0 になることはないと推測される。

第5章 考察

本研究で得られた推定結果について考察するため、流量調整槽①、②の設定温度を調整した際に、オペレータが期待しているガラス流量への効果（以下、オペレータ感覚）と、物理公式（A1 参照）を用いてガラス流量への効果を推定した結果とを比較した。表 5.1 に因果効果をまとめた結果を示す。

表 5.1 各方法における流量調整槽①、②の設定温度からガラス流量への因果効果まとめ

流量調整槽	設定温度を+1℃した時のガラス流量への効果 (kg / h)		
	統計的手法による推定	オペレータ感覚	物理公式による推定
①	0.7 ~ 3.2	30 ~ 50	7
②	0.3 ~ 2.7	10 ~ 30	10

まず、統計的手法による推定結果とオペレータ感覚を比較する。本研究で得られた統計的手法による推定結果では、流量調整槽①のほうが流量調整槽②よりも効果が大きい傾向だった。この傾向はオペレータ感覚と一致した。その一方で、効果の大きさはオペレータ感覚に比べ、統計的手法による推定結果のほうが小さいことがわかった。

安定生産期間中、オペレータは 0.1℃単位で設定温度を調整している。そのため、本分析の統計的手法による推定結果が正しいとすると、オペレータの介入がガラス流量へ与えている効果は最大でも 0.3 kg/h となる。安定生産期間のガラス流量の変動は約 10 kg/h はあることを考慮すると、0.1℃単位の流量調整槽の設定温度調整がガラス流量へ与えられる効果は非常に小さいといえる。本分析の結果から、オペレータの介入はガラス流量へほとんど影響を与えていない可能性が示唆された。

このような結果となった理由はいくつか考えられる。まず、オペレータの調整技量の差が影響したことが挙げられる。オペレータ感覚を調べるためにヒアリングを行った際、オペレータの中には、ガラス流量に与えている効果の大きさについて意識していない人もいた。そのため、流量調整槽の設定温度を調整する時にガラス流量にどれだけ効果を与えているか意識しているオペレータは、オペレータの中でも調整技量が高いと考えられる。そのように仮定した場合、オペレータ感覚の結果は、高い調整技量を持ったオペレータのみの結果と考えられる。その一方で、本分析で使用したデータセットの期間内において、流量調整槽の設定温度を調整したオペレータの調整技量は全員同じと仮定されている。このことから、オペレータ感覚のほうが、調整技量が高いオペレータを有意に抽出していたため、統計的手法による推定結果よりも効果が大きくなったことが考えられる。他には、オペレータが因果効果を誤認している可能性が考えられる。安定生産期間ではなく、大きくガラス流量が変動してしまった、いわゆるトラブル発生時において、オペレータが 1~2℃単位での調整を実施す

る場合がある。そのような経験から、トラブル発生時の効果を安定生産期間中の介入と同一視し、効果を誤認した可能性が考えられる。最後に、人的介入が効果的であるという認識バイアスによる効果の過大評価が可能性として挙げられる。

次に、物理公式による推定結果について考察する。物理公式を使用して推定された流量調整槽①、②の設定温度からガラス流量への効果は、流量調整槽①設定温度の因果効果のほう、流量調整槽②設定温度の因果効果よりも小さかった。これは、本研究で得られた統計的手法による推定結果、およびオペレータ感覚のどちらとも傾向が一致しなかった。この物理公式による推定では、設備設計時にガラス流量への因果効果が決まると考えられる。しかし、オペレータは調整時に設備に流れている電流値の大きさによって、ガラス流量へ与えられる効果の大きさは異なるという感覚を持っている。物理公式による推定では、設備に流れている電流値が変数として考慮されないため、このような差が生じたと考えられる。また、物理公式による推定結果よりも、統計的手法による推定結果のほうが因果効果は小さかった。物理公式では、設定温度を+1℃した時、その増加分の熱がすべてガラスへ伝わると仮定している。しかし、実際には大気中に放熱して失われる分も存在すると考えられる。この放熱した影響を正確に推測することは困難であるため、物理公式による推定では、放熱する分は考慮していない。その一方で、統計的手法による推定結果では、観察データを用いているため、放熱した影響も考慮して推定されていると考えられる。このことから、統計的手法による推定結果のほうが、因果効果が小さく推定されたと考えられる。

いずれの手法においてもガラス流量への因果効果を完全に推定したと断言することは難しい。しかし、統計的手法による推定結果は、個人の勘や経験によるバイアスがかかっている可能性が否定できないオペレータ感覚に比べ、客観性に秀でていると考えられる。また、統計的手法による推定結果は、誤差を含めて推定が可能な分、物理公式を用いた推定結果よりもオペレータ感覚に近いと考えている。

第6章 結論

6.1 本研究のまとめ

本研究では、因果推論の構造方程式モデルを適用し、オペレータが介入している流量調整槽の設定温度調整がガラス流量に与えている影響の評価を試みた。分析の結果、安定生産期間において、オペレータ感覚より統計的手法によって推定された因果効果は小さいことがわかった。この結果から、安定生産期間においてオペレータが介入している流量調整槽の設定温度調整は、ガラス流量へほとんど影響を与えられていない可能性が示唆された。

6.2 今後の課題

今後の課題は、本研究の推定結果を用い実証実験を行うことである。具体的には、推定結果として得られた値から、オペレータが期待している $3\sim 5\text{ kg/h}$ の調整を行うためには、安定生産期間でのオペレータの調整を $0.1\text{ }^{\circ}\text{C}$ 単位から $1\sim 2\text{ }^{\circ}\text{C}$ 単位へ変更する必要があることが推定された。または、オペレータの調整はガラス流量へほとんど影響を与えられていないことが示唆されたため、流量調整槽の設定温度を全く変更せずに、ガラス流量がどの程度変動するのか記録することも実験案として挙げられる。オペレータが調整せずともガラス流量の変動がこれまでの水準と変わらないのであれば、オペレータの流量調整槽の設定温度調整の時間を他の作業のために使用できるため、作業負担を減らせると考えられる。

これらの実証実験を行う場合は、トラブルが発生したとしても緊急停止が行えるような定修工事の前など、タイミングを見極める必要がある。その結果を踏まえて、よりガラス流量を安定化するための調整する方針や調整方法の検討を行う助けになると考えている。

最後に、今回使用した因果推論の考え方を異なる製造工程に適用し評価することも今後の課題として検討している。

謝辞

本論文の執筆にあたり、多くの方々にご支援していただきました。

基本的な統計学の考え方や研究の進め方、細かい言葉の表現など、様々なことを丁寧にご指導してくださった清水先生に心より感謝申し上げます。統計的因果推論と統計的因果探索の違いもわかっていなかった私に、基本的なことも含めて言葉を変えて説明してくださったり、同じようなことを何度聞いても嫌な顔一つせず教えてくださったりと、私にとって相談しやすい環境を整えてくださって本当にありがとうございました。

清水ゼミの皆様にも感謝申し上げます。色々な研究分野やテーマに触れることができ有意義な時間を過ごすことができました。ありがとうございました。一緒に学ぶことができた大学院同期にも感謝申し上げます。授業の課題やコンペと一緒に取り組んだ経験や、お昼ご飯や飲み会で色々な話をした思い出は私にとってかけがえのない財産です。他にも、輪読会で一緒した方々とは本の内容だけでなく、そこから派生した内容についても学ぶことができて、景色が広がりました。ありがとうございました。

滋賀大学大学院の先生方にも、感謝申し上げます。知識が不十分な私が考えた研究テーマに対して多くの先生方から真摯な助言をいただくことができて、とても貴重な経験を積むことができました。また、この派遣社会人制度を紹介してくださった笛田先生、中途半端な知識しか持たなかった私が、このような恵まれた環境で統計学や機械学習に関する体系的な知識を得るきっかけをくださって、本当にありがとうございました。

ここからは、会社の方々に向けてお伝えしたいと思います。人事部の井上さん、毎月資料や報告書にコメントくださり、大変励みになりました。ありがとうございました。ディスプレイ事業部の皆様にも、心より感謝申し上げます。毎日の朝会に出席し、ディスプレイ事業部に関わる事柄についてより多くの知識を得ることができました。中村部長、玉村課長には、私の分析に対していつも温かいコメントをくださり、大変励みになりました。また、橋本さんには、私の基本的な質問にも丁寧に答えてくださったり、因果グラフ作成の相談で多くのお時間をいただいたりと、大変お世話になりました。本当にありがとうございました。

毎月の進捗報告会にご参加いただいた工務部の常岡 GL、森保さん、東さん、前田さんにも感謝申し上げます。業務経験豊富な皆様からコード実装に関する助言をいただいたり、分析結果についてのコメント、質問など、様々な議論をすることができて、非常に心強かったです。本当にありがとうございました。

そして、製造技術部の高谷部長、東條 GL、メンタルを崩しかけて、もうできないと弱音を吐いた私に温かい言葉をかけてくださった心遣いに、言葉をどれほど尽くしても足りないほど感謝しております。自分自身の情けなさをどうにか受け入れてここまで研究を進めることができたのは、お二人のお陰です。心より感謝申し上げます。

最後に、大学院に進学することを快諾し、支えてくれた家族に心から感謝します。

Appendix

A1 ハーゲン・ポワズイユの式

配管内を流れる流量 $Q[kg/s]$ と圧力損失 $\Delta p[Pa]$ の関係式を表した流体力学の公式である。半径 $R[m]$ 、長さ $L[m]$ の配管内に、密度 $\rho[kg/m^3]$ 、粘性係数 $\mu[Pa \cdot s]$ の非圧縮性流体が層流定常流で流れていると仮定した時、式 (A1.1) が成り立つ。

$$\Delta p = \frac{8\mu L Q}{\pi R^4} \quad (A1.1)$$

この公式から、圧力損失に最も強く影響を与えているのは、4 乗がかかっている内径 R の逆数であることがわかる。図 3.1 より、流量調整槽②のほうが流量調整槽①よりも内径 R は小さいため、同じ流量 Q を流す場合、内径 R が小さい流量調整槽②のほうが圧力損失は大きくなる。この発生した圧力損失を補うために、流量調整槽②がより多く流量 Q が流れると推定される。

A2 データセットを変更した分析

本論では探索的に分析を実施したため、表 A2.1 に示す新たな 3 つのデータを用いて本論と同様の流れで分析を行い、因果効果の傾向を調べた。データセット①は本論で分析に使用したデータセットで、データセット②～④が新たに分析するデータセットである。

表 A2.1 使用したデータセットの期間と観測数一覧（データ取得頻度 1 分）

データセット	期間				観測数
①	2023/11/17	6:22	～	2023/11/30 3:53	18,572
②	2023/12/05	15:42	～	2023/12/12 0:00	9,139
③	2023/12/21	16:07	～	2023/12/28 23:10	10,504
④	2024/03/07	11:20	～	2024/03/14 13:38	10,219

データセット②～④を分析して得られた流量調整槽①設定温度からガラス流量への因果効果の推定結果を表 A2.2 に示す。

表 A2.2 データセット②～④を使用した
流量調整槽①設定温度からガラス流量への因果効果の推定結果まとめ

Lag	データセット②					データセット③					データセット④				
	p	β	S.E.	調整済み R^2		p	β	S.E.	調整済み R^2		p	β	S.E.	調整済み R^2	
7	2	1.33	0.82	0.22		3	0.33	0.47	0.19		1	1.56	0.48 *	0.18	
8	2	1.02	0.82	0.21		3	1.03	0.47	0.19		1	0.45	0.48	0.16	
9	1	0.40	0.83	0.19		2	1.40	0.48 *	0.17		1	0.16	0.48	0.16	
10	1	2.03	0.84	0.18		1	1.34	0.48 *	0.16		1	1.03	0.48	0.15	
11	2	1.84	0.83	0.20		1	1.90	0.48 *	0.13		2	1.90	0.48 *	0.18	
12	1	0.78	0.84	0.18		1	2.34	0.48 *	0.14		1	2.58	0.48 *	0.16	
13	2	1.73	0.82	0.22		1	2.04	0.49 *	0.13		1	1.69	0.48 *	0.15	
14	1	2.69	0.83 *	0.20		1	1.69	0.48 *	0.13		1	1.36	0.48 *	0.15	
15	1	1.73	0.84	0.19		1	1.64	0.49 *	0.13		1	1.46	0.49 *	0.14	
16	2	3.27	0.83 *	0.20		1	1.67	0.49 *	0.13		1	0.86	0.49	0.14	

Note. * <0.005

使用したデータセット、および Lag によって p は異なる結果となった。データセット①で得られた結果と比較すると、データセット②～④を使用した場合、p は小さい値をとる傾向が見られた。これは、BIC がサンプルサイズに影響を受けるためと考えられる（Lorach & Womack 2019）。

また、使用したデータセットによって流量調整槽①設定温度からガラス流量への因果効果の最大値と、最大値をとる Lag は異なった。これは、データセットによってオペレータの人員編成の変化などが影響している可能性が考えられる。

データセット①～④すべてで因果効果が有意であるといえたのは、流量調整槽①設定温度の調整から 14 分後だった。また、データセット①、③、④の結果のみを考慮すると、流量調整槽①設定温度の調整から 11～15 分後までで因果効果が有意であることがわかった。

次に、流量調整槽②設定温度について同様の分析を実施し、得られた推定結果を表 A2.3 に示す。

表 A2.3 データセット②～④を使用した
流量調整槽②設定温度からガラス流量への因果効果の推定結果まとめ

Lag	データセット②					データセット③					データセット④				
	p	β	S.E.	調整済み R^2		p	β	S.E.	調整済み R^2		p	β	S.E.	調整済み R^2	
7	2	2.33	0.53 *	0.23		3	1.80	0.32 *	0.19		1	1.55	0.38 *	0.18	
8	2	0.87	0.54	0.21		3	1.67	0.32 *	0.19		1	1.84	0.39 *	0.16	
9	1	0.49	0.54	0.19		1	1.47	0.32 *	0.14		1	1.31	0.39 *	0.16	
10	1	0.74	0.55	0.18		1	1.46	0.32 *	0.15		1	1.01	0.39	0.15	
11	2	1.18	0.54	0.20		1	1.18	0.32 *	0.13		2	1.42	0.39 *	0.17	
12	1	1.17	0.54	0.18		1	1.36	0.32 *	0.14		1	0.35	0.39	0.16	
13	2	0.99	0.53	0.22		1	1.21	0.32 *	0.12		1	0.53	0.39	0.15	
14	1	1.01	0.54	0.20		1	0.97	0.32 *	0.13		1	0.50	0.39	0.15	
15	1	1.67	0.54 *	0.19		1	0.96	0.32 *	0.12		1	0.53	0.39	0.14	
16	2	1.14	0.54	0.20		1	1.34	0.32 *	0.12		1	1.13	0.39 *	0.14	

Note. * < 0.005

流量調整槽②設定温度においては、データセット①～④すべてで統計的に有意であるといえたのは調整から 7 分後だった。また、データセット①、③、④の結果では、調整から 7～9 分後に統計的に有意であるといえる。p の値や最大の因果効果については、流量調整槽①設定温度と同じ傾向を示した。

これらの結果をふまえて、流量調整槽①設定温度の調整から 11～15 分後、流量調整槽②設定温度から 7～9 分後にガラス流量への因果効果があることが示唆された。また、いずれのガラス流量への因果効果の値でも、本論と同様に、オペレータが期待しているよりもその値は非常に小さかった。

A3 仮定を緩めた分析

A3.1 モデルの線形性とオペレータの調整技量

本論では線形モデルを仮定し分析を行ったが、この仮定が満たされていない場合を想定し、追加で分析を行った。まず、使用するモデル、使用するデータによって想定される仮定をまとめた概要図を図 A3.1 に示す。

線形モデルを仮定した場合、ガラス流量と流量調整槽の設定温度は直線の関係になる。しかし、ガラス流量と流量調整槽の設定温度は正確には指数関数の関係にある。そのため、非線形モデルを使用し同様に分析することで、本論で線形モデルの仮定を置いた妥当性を確認することとした。

また、本論での分析は長期間データを使用しており、複数人のオペレータの平均値が推定結果として得られている。これはオペレータの調整技量は全員同じという仮定が置かれていることと同じである。しかし、オペレータの調整技量は人によって異なると言われている。そこで、1人のオペレータだけが調整している1勤務分のデータセット（以下、短期間データ）を使用し、オペレータの調整技量が全員同じという仮定を緩めた分析も合わせて行うこととした。

なお、これ以降の分析は流量調整槽①設定温度のみの結果を示す。

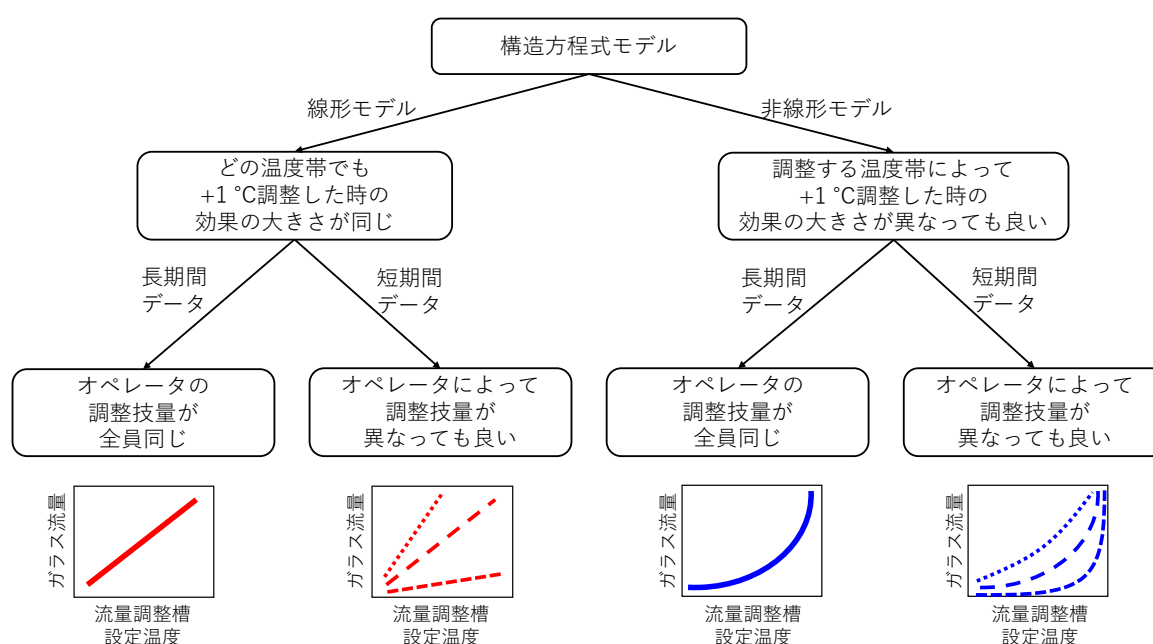


図 A3.1 使用モデルと使用データごとに想定される仮定、およびガラス流量と流量調整槽設定温度との関係性のまとめ

A3.2 短期間データを用いた線形モデル分析

ここでは、線形モデルを仮定し、短期間データを用いた分析結果について示す。データセット①の結果を踏まえて、Lag=13 に固定し、オペレータ K、L、M の 3 人の 1 勤務分の短期間データを使用して、本論と同様の手順で分析を行った。1 勤務分の観測数はおよそ 500 である。

表 A3.1 に短期間データを用いた線形モデルによる流量調整槽①設定温度からガラス流量への因果効果推定結果を示す。

表 A3.1 短期間データを用いた線形モデルによる
流量調整槽①設定温度からガラス流量への因果効果推定結果のまとめ

オペレータ	p	β	S. E.	95%信頼区間		調整済み R^2
				Lower	Upper	
K	1	4.87	3.04	-1.10	10.85	0.23
K	1	1.76	2.73	-3.62	7.14	0.25
K	1	-0.57	3.12	-6.69	5.57	0.27
K	1	2.20	3.95	-5.56	9.96	0.28
K	1	5.63	2.72 *	0.29	10.97	0.24
L	1	< 0.01	2.77	-5.43	5.44	0.15
L	1	-0.57	3.12	-6.69	5.57	0.27
L	1	0.31	2.89	-5.38	5.99	0.33
M	1	0.42	2.44	-4.39	5.23	0.16
M	1	2.97	5.72	-8.27	14.21	0.26
M	1	5.17	2.95	-0.62	10.96	0.27
M	1	-1.35	3.11	-7.47	4.77	0.22
M	1	-3.92	4.41	-12.58	4.75	0.25

Note. * <0.05

分析した結果、オペレータ K の 1 つの期間のみ統計的に有意であるといえたが、それ以外の期間では統計的に有意でないことがわかった。統計的に有意であった期間での因果効果は 5.63 と、長期間データで得られた因果効果、2.44 に比べて 2 倍以上であった。このことから、オペレータによって調整技量に差があることが示唆された。

ほとんどの期間で統計的に有意でない結果となった理由としては、ガラス流量は誤差が大きい、弱いデータであったことから、サンプルサイズを小さくした影響を受けたためと考えている。

A3.3 長期間データを用いた非線形モデル分析

ここでは、非線形モデルを仮定し、長期間データとしてデータセット①を用いた分析結果について示す。非線形モデルとしては、人の思考に近いと考えられる Random Forest を使用することとした。

A2 の結果をふまえ、Lag=11~15 の各時点のみ分析することとした。まず、Python の scikit-learn ライブラリーを使用し、n_estimators=100 を指定して、データセット①のデータを用いてガラス流量を予測するモデルを作成した。流量調整槽①設定温度として介入させた値を指定し（例えば 1202.2 °C で流量調整槽①設定温度を固定し）、親変数はデータセット①と同じ値を用いて、上記で作成したモデルによりガラス流量の予測値（以下、予測ガラス流量）を得た。得られたガラス流量の分布をガウス分布として見なせると考え、得られた予測ガラス流量の平均値を予測ガラス流量の期待値として算出した。データセット①では、流量調整槽①設定温度は 1202.2 °C から 1205.9 °C までの範囲で調整されていたため、その範囲内で 0.1 刻みの値を介入値として用い、それぞれの流量調整槽①設定温度から予測ガラス流量の期待値を算出した。

x 軸に介入した流量調整槽①設定温度、y 軸に予測ガラス流量の期待値からガラス流量の目標値を引いた値のプロットの結果を、Lag ごとに図 A3.2 に示す。なお、ガラス流量の値は機密性が高いため、このような処置をとった。

いずれの Lag でも、A3.1.1 の概要図のような下に凸のグラフにはならなかった。細かくみるとグラフには凹凸があるが、グラフのおおよその形から、線形モデルとして近似しても問題ないように考えられる。また、このグラフの傾きから算出できる因果効果は、流量調整槽①設定温度を +1 °C した時、ガラス流量への効果は最大でも 0.02 kg/h と、オペレータが期待している効果より非常に小さかった。

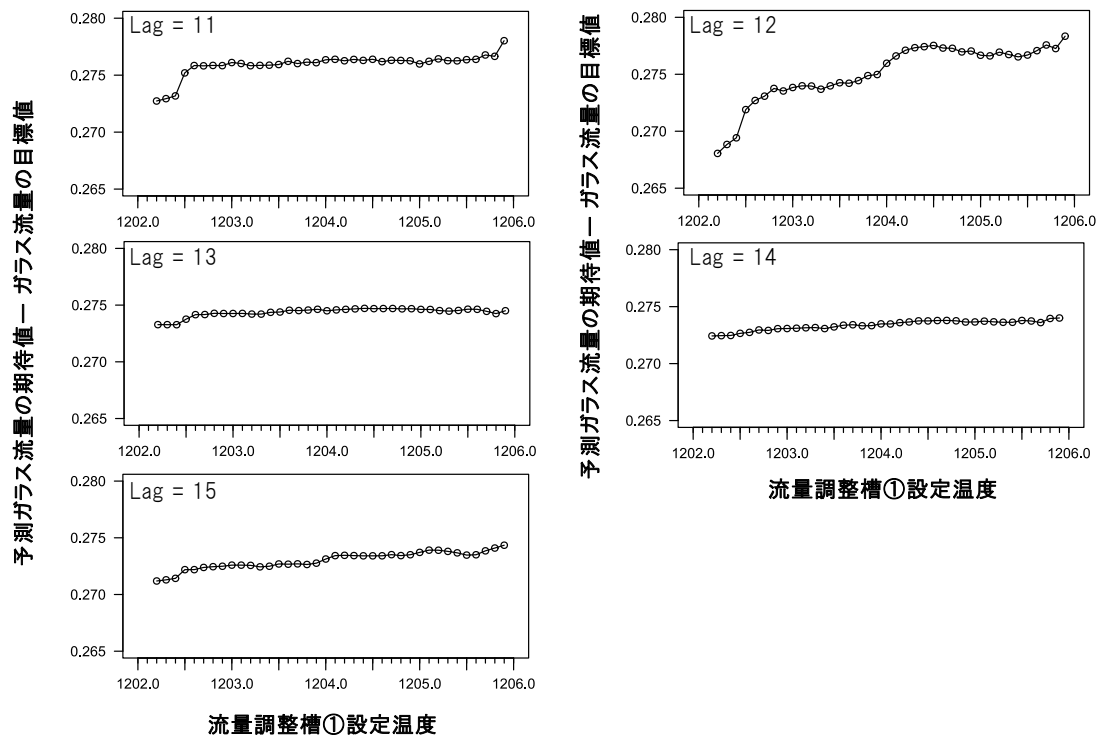


図 A3.2 長期間データを用いた非線形モデル（Random Forest）による
 流量調整槽①設定温度の介入値と
 予測ガラス流量の期待値とのガラス流量の目標値との差分プロット

A3.4 短期間データを用いた非線形モデル分析

ここでは、非線形モデルを仮定し、短期間データを使用した分析結果を示す。A3.2 の短期間データ線形モデルの分析結果をふまえて、統計的に有意であるといえたオペレータ K のデータを用い、A3.3 と同様にモデルを作成して予測ガラス流量の期待値を得た。結果を図 A3.3 に示す。

1 勤務分のデータでは、流量調整槽①設定温度を変化させた範囲が 0.5°C と小さく、グラフの形を確かめるには不十分だった。

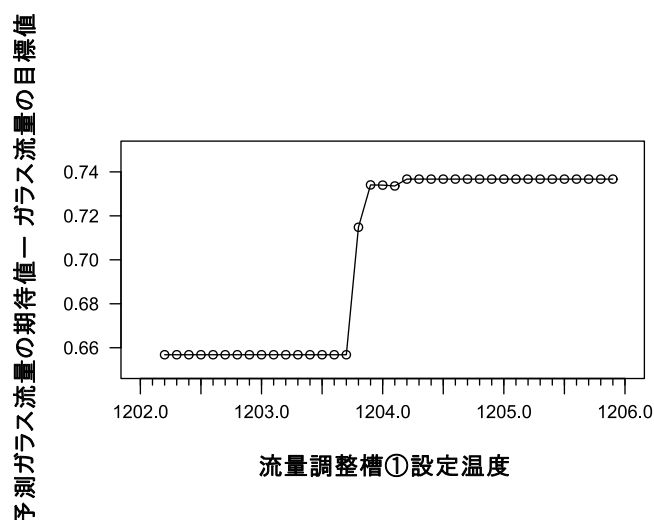


図 A3.3 短期間データを用いた非線形モデル (Random Forest) による
流量調整槽①設定温度の介入値と
予測ガラス流量の期待値とのガラス流量の目標値との差分プロット

参考文献

- 山根正之, 安井至, 和田正道, 国分可紀, 寺井良平, 近藤敬, 小川晋永. ガラス工学ハンドブック. 朝倉書店, 1999.
- 日本電気硝子株式会社. 板ガラス製造方法、清澄容器及び板ガラス製造装置. 特開 2018-95535. 2018-06-21.
- 松岡純. ガラス融液の物性. *NEW GLASS*, 120, Vol.32, No.1, pp.4-9, 2017.
- Lennart Ljung. System Identification: Theory for the User 2nd Edition, *PTR Prentice Hall Information and System Science Series*, 1998.
- 足立修一. ユーザーのためのシステム同定理論. 社団法人 計測自動制御学会, 1993.
- Swell Wright. Correlation and causation. *Journal of agricultural research*, Vol. 20, No. 7, pp. 557-585, 1921.
- Herman Wold. Causality and econometrics. *Econometrica: Journal of the Econometric Society*, pp. 162-177, 1954.
- Shohei Shimizu, Patrik O Hoyer, Aapo Hyvärinen, Antti Kerminen, and Michael Jordan. A linear non-gaussian acyclic model for causal discovery. *Journal of Machine Learning Research*, Vol. 7, No. 10, 2006.
- 黒木学(工学). 構造的因果モデルの基礎. 共立出版, 2017.
- Carlo E. Bonferroni. Teoria statistica delle classi e calcolo delle probabilità. *Pubblicazioni del Istituto Superiore di Scienze Economiche e Commerciali di Firenze*, 8, 1-62.
- Carlos Cinelli and Chad Hazlett. Making sense of sensitivity: Extending omitted variable bias. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, Wiley Online Library, 82, pp. 39-67, 2020.
- Julie Loarch, Andrew Womack. Value of sample size for computation of the Bayesian information criterion (BIC) in multilevel modeling. *Behavior Research Methods*, 51, pp. 440-450, 2019.