

令和 2 年 5 月 29 日現在

機関番号：82401

研究種目：挑戦的研究（萌芽）

研究期間：2018～2019

課題番号：18K19033

研究課題名（和文）時系列データの機械学習による結晶成長の非再現性の追求

研究課題名（英文）Machine learning study on non-reproducibility of crystal growth results

研究代表者

沓掛 健太郎（Kutsukake, Kentaro）

国立研究開発法人理化学研究所・革新知能統合研究センター・研究員

研究者番号：00463795

交付決定額（研究期間全体）：（直接経費） 4,900,000円

研究成果の概要（和文）：結晶成長では、同一の条件（装置、試料、成長レシピなど）にもかかわらず、結果が異なることがある。本研究では機械学習を活用して、非再現性の要因を特定することを目指した。結晶成長装置に多数のセンサを設置し、データを収集・一元管理するマルチセンシングシステムを設計・構築した。さらに得られた時系列データをリカレントニューラルネットワークを用いて機械学習することで、将来の温度変化をリアルタイムで予測し、また各パラメータの影響を定量評価することに成功した。これらの結果は、より高精度な結晶成長制御につながる成果である。

研究成果の学術的意義や社会的意義

本研究は、シミュレーションなどでは考慮することが難しい装置内部の状態の微妙な変化を、実実験でのマルチセンシングによって検出し、その影響を機械学習によって定量化することを目指したものである。実実験におけるデータ取得の指針や時系列データの機械学習の検討など、当初の目的である非再現性の要因追及を越えて、この分野の発展の基礎となる多くの成果が得られた。

研究成果の概要（英文）：In crystal growth, the results are often different even though under the same conditions, such as the same a furnace, material, growth recipe, etc. In this research, we aimed to identify the factors of this non-reproducibility using machine learning. We designed and constructed a multi-sensing system with several sensors in a crystal growth furnace, and collected and managed the data. Furthermore, we succeeded in predicting future temperature changes in real time and quantitatively evaluating the influence of each parameter using a recurrent neural network. These results lead to more accurate crystal growth control.

研究分野：応用物理、結晶工学、結晶成長、機械学習

キーワード：結晶成長 機械学習 実験再現性 リアルタイム予測 結晶工学 その場観察 マルチセンシング

## 様式 C - 19、F - 19 - 1、Z - 19 (共通)

### 1. 研究開始当初の背景

結晶成長では、同一の条件(装置、試料、成長レシピなど)で実験を行なったにもかかわらず、核形成タイミング、成長速度、結晶品質といった結晶成長の結果が異なることがしばしば起こる。今日、様々な結晶で成長の条件やプロセスは複雑化・極限化する傾向にあり、よりシビアな条件では非再現性の問題はますます深刻となる。

非再現性の解明に向けた一つのアプローチは、シミュレーションの活用である。結晶成長装置を計算機内で正確に再現することで、内部で起こっている複雑な現象の理解や各パラメータの影響解析を通して、非再現性の原因を明らかにすることができる。しかし一方で、装置内の部材の位置や物性、装置内外の環境の微妙な変化が、非再現性に影響を及ぼしていることも考えられる。このような微妙な変化を一つ一つ検出して、さらにシミュレーションに正確に反映することは現実的には難しいことが多い。

非再現性の解明に対する別のアプローチは、データから帰納的に要因を特定することである。ここでは機械学習が有効である。機械学習は、大量データをもとに、コンピュータを用いて有用な法則を求める方法であり、複雑な因果関係に対しても帰納的に関係を求めることができる。自動運転のための物体判別や進路予測、小売業の売り上げ予測、自動翻訳など、様々な分野で活用が進んでいる。特に材料研究分野では、機械学習を利用したマテリアルズインフォマティクスが大きな盛り上がりを見せており、材料の結晶構造(X)と特性(Y)の関係に基づいた新材料探索などで目覚ましい成果をあげている。本研究の対象である結晶成長の場合は、成長条件(X)と実験結果(Y)を機械学習を用いて関係づけることで、実験結果の非再現性の要因となるXを特定することができると考えられる。しかし、従来のマテリアルズインフォマティクスの多くはシミュレーション結果を利用したものであり、実実験への応用は、機械学習のためのデータ取得の方法を含めて、発展途上である。そこで本研究では、結晶成長を例にとり、実実験において機械学習を用いて非再現性の要因を追求することを目指した。

### 2. 研究の目的

実実験の非再現性の機械学習を行うためにはまず、非再現性に影響を与える要因(X)の微妙な変化を検出することが必要である。そのために本研究では、結晶成長装置内に設置した複数のセンサを用いて、機械学習に必要なデータの収集と管理を行う「マルチセンシング」を研究項目とした。また、結晶成長で得られるデータは時間方向の変化を伴う時系列データである。しかし、結晶成長の時系列データに対して、どのような前処理および機械学習を行い、非再現性に関する情報を得るかについての系統的な研究はない。そこで本研究では、データの前処理を含めて「結晶成長時系列データの機械学習」を研究項目とした。これらの研究結果を合わせて、「結晶成長その場観察 × 大量データ」によって、各パラメータの影響の解明、すなわち非再現性の要因を特定することを目的とした。さらに、得られた機械学習モデルを活用して、将来の結晶成長装置内の温度変化をリアルタイムに予測することにも挑戦した。

### 3. 研究の方法

結晶成長実験は、東北大学金属材料研究所の融液成長その場観察装置(図1)を利用して行った。同装置では、結晶が成長する様子を顕微鏡によって直接観察し、動画像として取得することができる。結晶成長のモデル材料としてはシリコンを選択した。シリコンは結晶化潜熱が大きいので、潜熱による温度場の変化が大きい。そのため、ヒータ設定温度などその瞬間の情報のみでは結晶成長予測が難しく、結晶成長炉内の環境に加えて、それまでに温度がどのように変化してきたか、結晶がどのように成長してきたかといった多角的な情報が必要となる。そこで本研究では、ヒータ設定温度、ヒータ出力、炉内モニタ温度、気温、湿度、ガス流量、シャッター開閉などの多様な装置情報を取得し、それらを時系列データとして結晶成長のその場観察動画と時間同期させて管理する炉内環境モニタシステムを構築した。

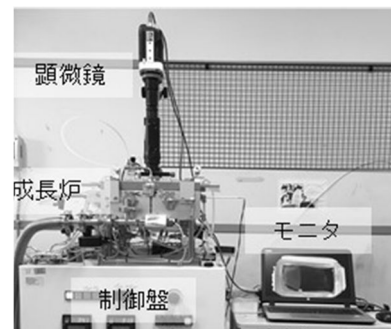


図1 融液成長その場観察装置

機械学習には、汎用の機械学習ライブラリである Keras を使い、リカレントニューラルネットワークによってモデルを作成した。ここで、機械学習のハイパーパラメータに加えて、入力するパラメータの種類、値の範囲(時間方向および絶対値範囲)、前処理方法(正規化、二値化、外れ値処理など)を検討した。

さらに、炉内環境モニタシステムからの出力を直接機械学習モデルに入力し、予測を実行することで、結晶成長装置の変化と同期して、将来の炉内温度分布変化の予測値をリアルタイムで表示させるシステムを構築した。

### 4. 研究成果

まず、既存の結晶成長その場観察装置に対して改造を施し、炉内温度、ガス流量、気温、湿度などの炉内状況に関する情報を多様なセンサ類によって取得する機構を設計・設置した。さらに、

これらの観測情報と、ヒータ出力、ヒータ設定温度などの装置情報および結晶成長のその場観察動画を時間同期させ、時系列データとして一元管理する結晶成長データ収集システムを設計・構築した。最終的には、30 超のパラメータについての情報を収集するシステムとなった。

次に、得られた時系列データの前処理を検討した。ここでは装置および結晶成長についての知識を活用することで、より意味のあるデータとすることを行った。例えば、のぞき窓のシャッタ開閉は、のぞき窓位置の連続値として得られるが、結晶成長への影響を考えると、開もしくは閉の二値で影響する。そこで、閾値によって開閉の二値へと変換し、機械学習に用いた。またデータを詳細に解析した結果、今回の実験データでは、シリコン融点より約 400 以上低い温度範囲では、急冷や低温保持など、目的とする結晶成長とは異なる種類の温度変化のデータも含んでいた。そこで、機械学習に用いるデータは融点に近い高温範囲に限定した。このような結晶成長実験の内容に沿った前処理を行うことによって、機械学習の予測精度を向上することができた。

以上の結晶成長実験データを用いて、リカレントニューラルネットワークによって、過去 60 秒間の時系列データおよび未来のヒータ出力設定プログラムから、未来 30 秒後までの炉内温度変化を予測する機械学習モデルを作成した。得られた機械学習モデルのテストデータ（学習に用いていないデータ）に対する予測値と実測値の関係を図 2 に示す。平均二乗誤差 (RMSE) および R2 スコアはそれぞれ、1.0、0.99 であり、実用可能な精度を得ることができた。

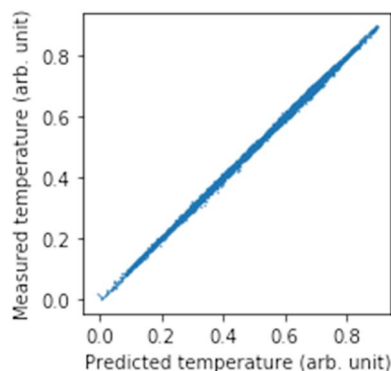


図 2 成長炉内温度の測定値と予測値の関係

さらに、得られたモデルを用いて、各パラメータが未来の温度へ与える影響を調べた。図 3 にのぞき窓シャッタ開閉が温度変化に与える影響の予測を示す。ある結晶成長実験の試料温度変化について、時刻 0 まではシャッタ閉状態での測定結果を示し、時刻 0 以降はシャッタが開閉それぞれの状態に対する温度変化の予測である。シャッタを開けることによって窓からの抜熱が発生し、温度上昇が緩やかになることがわかる。この結果が示すように、シャッタ開が温度上昇を抑制する効果を定量的に見積もることができた。このシャッタ開閉効果の予測は、実験を行った範囲内でのあらゆるヒータ出力・温度の組合せに対して行うことができるため、多様な実験条件の組合せに対応することが可能である。

さらに、マルチセンシングシステムを介した結晶成長装置からの出力をパラメータ変換し、機械学習モデルに直接入力することで、未来の温度変化をリアルタイムで予測するシステムを構築した。機械学習は、学習には比較的長い計算時間を要するが、予測は計算量が少ないため瞬時に値を出力することができる。そのため、本研究で構築したリアルタイム予測システムは、結晶成長装置における実際の操作に追従して、瞬時に予測結果を出力することが可能である。将来の温度変化に関するこのような情報は、特に、結晶が成長する様子を観察しながらマニュアル操作によって結晶育成を行う場合に、操作の指針を与えるものとして非常に役に立つ。

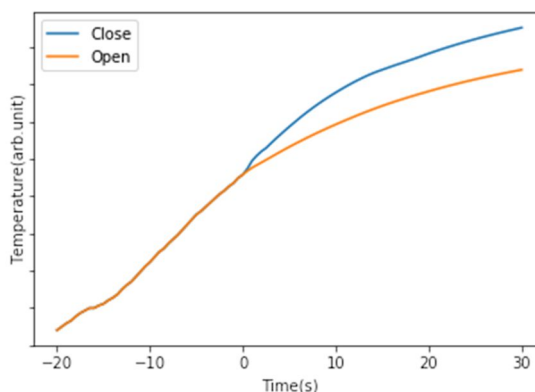


図 3 成長炉内温度の予測。時刻 0 以降の温度は、のぞき窓シャッタの開閉に対応。

以上のように、機械学習を用いた炉内環境の時系列変化が将来の環境変化に与える影響の定量は、結晶成長の非再現性の要因解明に直結する有用な方法である。さらに、機械学習の高速予測を活かしたリアルタイムの将来予測は、これまでにない情報を与えることができ、新しい結晶成長制御につながると期待される。将来的には、機械学習を活用して、その時点までの情報を反映したリアルタイムな未来状態予測に基づいた結晶成長制御に取り組み、より高精度、高均一な結晶成長技術を目指す。

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計6件（うち招待講演 4件 / うち国際学会 2件）

1. 発表者名 沓掛健太郎、Boucetta Abderahmane、工藤博章、松本哲也、宇佐美徳隆
2. 発表標題 データ科学手法による結晶成長炉内の最適温度測定位置の検討
3. 学会等名 第47回結晶成長国内会議
4. 発表年 2018年

1. 発表者名 Kentaro Kutsukake, Yusuke Hayama, Tetsuya Matsumoto, Hiroaki Kudo, Tatsuya Yokoi, Yutaka Ohno, and Noritaka Usami
2. 発表標題 Generation and propagation of dislocations in multicrystalline silicon for solar cells
3. 学会等名 International Symposium on Modeling of Crystal Growth Processes and Devices (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 沓掛健太郎
2. 発表標題 結晶成長・結晶評価へのデータ科学活用
3. 学会等名 第151回結晶工学分科会研究会 いまからはじめるインフォマティクス~チュートリアルから先端事例まで~ (招待講演)
4. 発表年 2019年

1. 発表者名 Kentaro Kutsukake
2. 発表標題 Application of data science techniques to crystalline silicon research for solar cells
3. 学会等名 29th Workshop on Crystalline Silicon Solar Cells & Modules: Materials and Processes (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 沓掛健太郎
2. 発表標題 "AI vs 人間" 良い結晶を創るのはどっち?
3. 学会等名 第42回 結晶成長討論会 (招待講演)
4. 発表年 2019年

1. 発表者名 沓掛健太郎、前田健作
2. 発表標題 機械学習による結晶成長炉内温度の予測
3. 学会等名 第48回結晶成長国内会議
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分 担 者	前田 健作  (Maeda Kensaku)  (40634564)	東北大学・金属材料研究所・助教    (11301)	