

新物質の合成条件を効率よく推薦する手法を開発

ポイント

- 理論計算で予測された新物質も、実際に合成するには合成研究者の勘と経験に基づく多くの試行錯誤を伴う合成実験が必要だった。
- 1000 件規模の並列合成実験データを活用して、新物質の合成成否を効率的に予測する推薦システムの開発に成功した。
- 合成成否だけでなく、さまざまな特性値に対しても適用可能であり、広範な材料開発への応用につながる事が期待される。

概要

京都大学大学院工学研究科 材料工学専攻の林博之 助教、世古敦人 同准教授、田中功 同教授らの研究チームは、1000 件規模の並列合成実験^{注1)} データを活用して、新物質を合成するための実験条件を効率的に探索できる推薦システム^{注2)} の開発に成功しました。

近年、計算材料学や人工知能 (AI)^{注3)} を活用することで未知の物質を予測する研究が世界的に活発になっています。しかし、予測された新物質を実際に合成するためには合成研究者による勘と経験に基づく多くの試行錯誤の実験が必要です。AI を適用するための実験結果のデータベースを構築することがボトルネックとなり、多くの新物質が「絵に描いた餅」状態だったのです。

本研究で開発した手法により、約 24 万通りの合成条件から、合成可能な条件を効率的に予測することに成功しました。マテリアルズ・インフォマティクス^{注4)} 分野に突破口をもたらすと期待されます。

本研究成果は、2019 年 12 月 14 日 (米国東部時間) に国際科学誌「Chemistry of Materials」のオンライン版で公開されました。

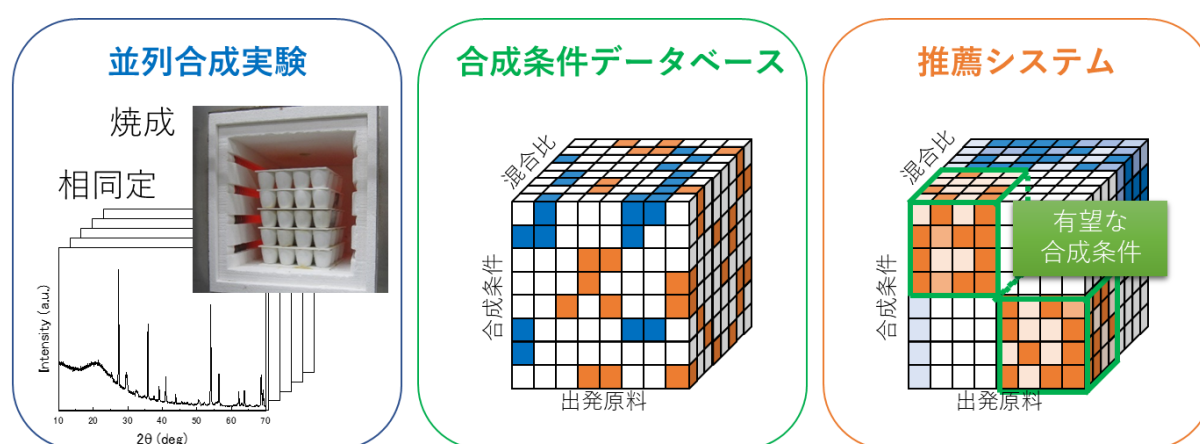


図 本研究の概念図

並列合成実験結果から作成した合成条件データベースに基づき、
未実験の合成条件から有望な条件を推薦する。

1. 背景

マテリアルズ・インフォマティクスでは、蓄積したデータをもとに、AI を活用して新物質・新材料の予測、合成、評価など一連の研究開発プロセスを効率化します。それには当然データが必要ですが、従来の材料研究はデータを系統的に蓄積することに重きを置いてこなかったため、人類が共有資産として使える材料データは多くありません。この状況を最近の5年ほどで変えたのが、量子力学に基づく第一原理計算^{注5)} 結果のデータベース構築です。この理論計算データに基づいて、多くの未知物質や未知特性の予測が行われ、材料研究は新しい局面を迎えました。

しかし、理論計算をもとに効率的に予測した物質や特性も、実証しなければ「絵に描いた餅」に過ぎません。実証実験の出発点は、予測された未知物質を実際に合成することですが、その工程にはいまだ AI を適用できておらず、合成研究者の勘と経験に基づく多くの試行錯誤の実験が必要でした。それは AI を適用するためのデータベースがなかったためです。実験結果を予測するためのデータベースは、断片的でない対象に対して一定のルールに従い系統的に行った実験による、均質なデータから構成されている必要があります。理論計算によるデータだけでは不十分なのです。

このような現状から脱却するため、本研究では、新物質の合成条件を AI で効率的に予測するためのシステム開発と、それに適した実験結果のデータベース構築を目指しました。

2. 研究手法・成果

まず、合成実験の成否に関するデータを系統的に得るために、無機化合物の並列合成実験システムを構築しました。一般に未知の無機化合物を合成する際には、合成方法、原料、焼成温度などさまざまな合成条件について、研究者が一つ一つ試料を合成しますが、並列合成実験では多数の試料を同時並行的に合成することができます。研究チームは、4種類の合成方法に対し、系統的に合成実験を行うシステムを開発しました。

対象としたのは2種類の陽イオンと酸化物イオンから構成される3元系酸化物です。23種類の原料を用い、そのさまざまな配合比に対して、5種類の焼成温度で、合計約1600通り合成しました。そして、目的の性質を持った物質ができたかどうかを、試料自動交換機を備えた粉末X線回折装置で連続的に評価、点数化し、約1600件の合成成否に関するデータベースを構築しました。図1に示すのはモリブデン酸アルミニウムを合成目的とした場合の例で、4種類の合成方法と5種類の焼成温度に対し、モリブデン酸アルミニウムが合成できているかどうかで点数を与えています。このような作業を、その他の3元系酸化物でも行います。

次に、この1600件の合成成否データを、探索対象となる約24万通りの全実験条件を表した合成条件テンソル（合成条件データの多次元配列）に入力しました。低ランク性の仮定^{注6)} のもとでテンソル分解することで、未実験の条件についての合成実験の成功可能性を予測しスコアとして与える推薦システムを開発しました（図2）。

得られた結果の妥当性を検証するため、交差検定をするとともに、高い成功可能性スコアを有する物質で実際に合成実験をしました。その結果、推薦システムが予想した通り、各物質でスコアの高い合成条件では合成に成功し、スコアが低い合成条件では合成が失敗することが分かりました。開発した手法で、新物質の合成条件の探索が大幅に効率化できることを示すものです。

3. 波及効果、今後の予定

本研究は、新物質の合成条件を AI で効率的に予測するためのシステム開発と、それに適した実験結果のデータベース構築について一定の成功を収めました。マテリアルズ・インフォマティクスに基づいた新物質の開

発に突破口をもたらすと期待されます。

マテリアルズ・インフォマティクス分野は世界的に研究者の参入が多く、年々飛躍的に前進しています。これまでのマテリアルズ・インフォマティクスは理論計算が主導して成果を上げてきましたが、これは前哨戦で、本丸は実験データを活用した研究開発の効率化だと研究チームは考えています。今後は、ロボットを用いた自動実験によるデータ収集と、AIによる解析、そして新たな実験計画というような流れで、研究開発が大きく加速すると期待されます。

本研究で構築したデータベースは、まだ1000件規模の小さなものですが、開発した手法は将来の自動実験を見据えており、容易に拡張できます。得られた物質の評価も、現在は粉末X線回折実験による合成成否判定だけですが、これをさまざまな特性値とつなげることで、開発した手法を適用できる材料の範囲も大きく広がると期待されます。

4. 研究プロジェクトについて

本成果は、以下の事業・研究領域・研究課題によって得られました。

戦略的創造研究推進事業 個人型研究（さきがけ）

研究領域：「理論・実験・計算科学とデータ科学が連携・融合した先進的マテリアルズインフォマティクスのための基盤技術の構築」

（研究総括：常行 真司 東京大学 大学院理学系研究科 教授）

研究課題名：「高効率な新物質発見のための合成手法推薦システムの構築」

研究者：林 博之（京都大学 大学院工学研究科 材料工学専攻 助教）

研究実施場所：京都大学 大学院工学研究科

研究期間：平成29年12月～令和3年3月

本研究領域では、実験科学、理論科学、計算科学、データ科学の連携、融合によって、それぞれの手法の強みを生かしつつ得られた知見を相互に活用しながら新物質・材料の設計に挑む先進的マテリアルズ・インフォマティクスの基盤構築と、それを牽引する将来の世界レベルの若手研究リーダーの輩出を目指しています。

<用語解説>

注1）並列合成実験

物質の合成にあたり、工程の条件をさまざまに変えて、系統的に実験を行う手法。より広い意味でコンビナトリアル実験と呼ばれる場合もある。

注2）推薦システム

ネットショップなどで、購買者の購入履歴データを活用して「おすすめ」を提示するために広く使われているAIシステム。

注3）人工知能（AI）

予測や解析など知的作業を人間に代わって計算機に行わせる技術。1950年代以降、研究が進められてきたが、最近では第3次AIブームと呼ばれ、研究熱が高まっている。

注4) マテリアルズ・インフォマティクス

材料データを AI に適用し、材料科学の諸問題の解決に役立てる技術。新材料の発見から工業化に至る一連の開発プロセスが大幅に効率化できると期待されている。

注5) 第一原理計算

量子力学の基本原理に基づいて物質の構造、安定性、物理的性質や挙動を計算する方法。実験値などの先験的情報を必要とせず、理論計算で予測できる。計算機と計算技術の進歩により、2000 年以降急速に材料科学への具体的な適用が進んだ。

注6) 低ランク性の仮定

与えられたテンソル／行列を小さなテンソル／行列と因子行列に分解しても、抽出したい情報に漏れが少ないと仮定すること。「おすすめ」を提示する推薦システムに使われている。

<研究者のコメント>

合成に成功した物質はその合成条件とともに論文として公開されていますが、その背後には研究者らの試行錯誤による膨大な数の合成失敗データがあり、それらは研究者各自のノートの中に眠っています。本手法は、失敗データであっても貴重なデータとして活用し、新しい物質の合成手法を見出すものです。このような状況は合成研究だけでなく様々な分野で共通した問題だと感じていますので、本研究が失敗データ活用の一例として研究開発のきっかけになればと思います。

<論文タイトルと著者>

タイトル：Recommender System of Successful Processing Conditions for New Compounds Based on a Parallel Experimental Dataset（並列合成データベースに基づいた推薦システムによる新規物質の合成成功条件予測）

著 者：Hiroyuki Hayashi, Katsuyuki Hayashi, Keita Kouzai, Atsuto Seko and Isao Tanaka

掲 載 誌：Chemistry of Materials DOI：10.1021/acs.chemmater.9b01799

* 本論文は「Just Accepted Manuscript」として 11 月 19 日に校正前の全文が公開されています。

< 参考図 >

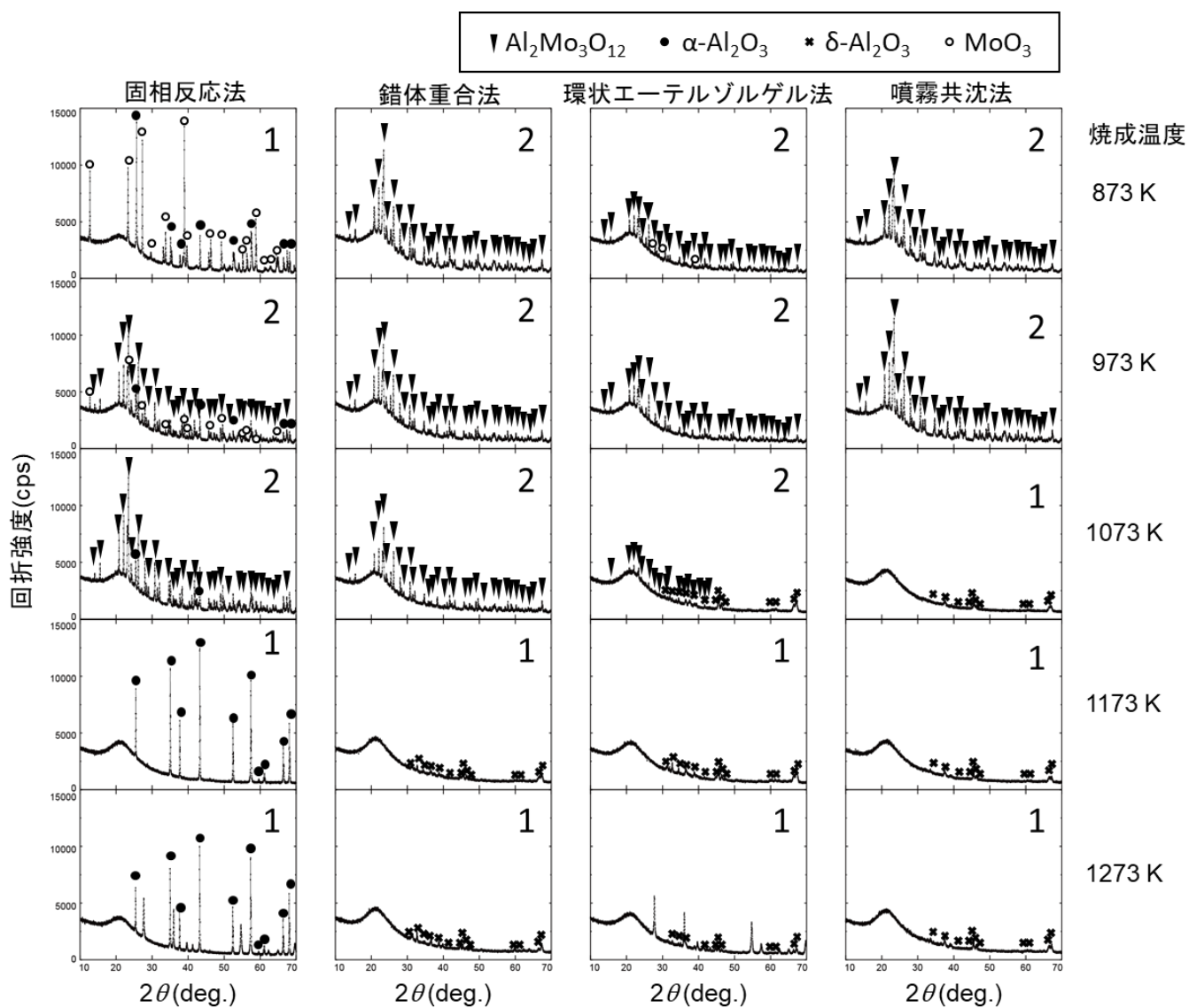


図1 並列合成実験における粉末X線回折実験結果

モリブデン酸アルミニウム ($\text{Al}_2\text{Mo}_3\text{O}_{12}$) の合成を目的として、4 種の実験方法で5 種の焼成温度で合成した粉末X 線回折実験結果。▼マークは目的物質に起因する回折ピークであり、目的物質が合成できた場合には2 点のスコアを与えた。一方、▼マークが見られない場合は合成に失敗したと判断し、1 点のスコアを与えた。

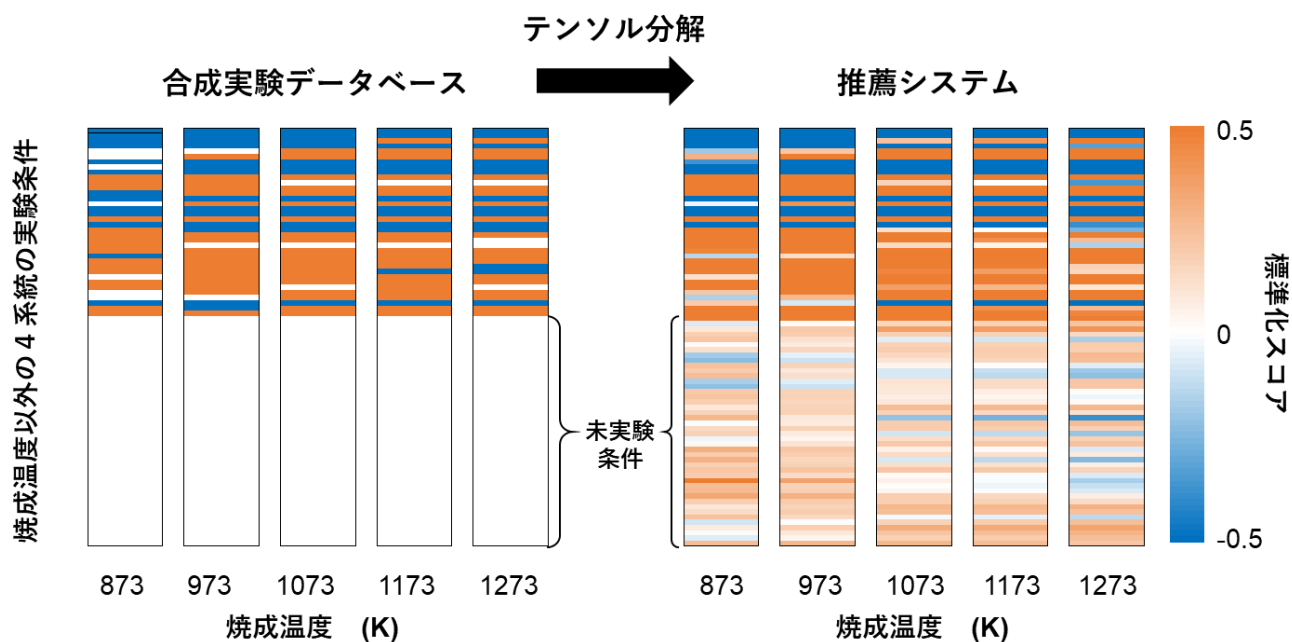


図2 テンソル分解による推薦システムの結果の例

テンソル分解による合成可能性スコアの変化。横軸は焼成温度を、縦軸は2種の原料と配合比および合成手法という4系統の実験条件を併せて示す。テンソル分解前の合成実験データベースでは、橙色が合成成功、青色が合成失敗、白色は未実験の条件を示す。テンソル分解後の推薦システムでは、未実験の条件に対しても成功可能性が標準化スコアとして予測されている。橙色が濃いほどスコアが高く、成功可能性が高いと期待できる。