
機械学習予測技術を用いたケミカルトナーの品質自動制御

Automatic Quality Control for Chemical Toner Using Machine Learning Prediction

高橋 直希* 宮川 将士* 佐藤 直之** 渡邊 好夫** 佐藤 敏明***
Naoki TAKAHASHI Masashi MIYAKAWA Naoyuki SATOH Yoshio WATANABE Toshiaki SATOH

加納 学****
Manabu KANO

要 旨

ポリエステル重合を用いたケミカルトナーの1種である重合トナーの品質制御は、過去の品質測定値を参照して品質管理者が工程の操作変数を決定する手動調整を実施していた。そのため、トナー品質管理者の多大な工数が必要であるとともに、品質安定性が不足していた。この問題を解決するため、品質予測システムと操作量算出システムで構成される品質自動制御システムを開発した。品質予測システムの構築では、いくつかの機械学習手法の中から、変数選択したRandom Forestモデルを選定した。操作量算出システムの構築では、品質予測値を参照して、操作量を自動算出するシステムを開発した。品質自動制御システムをトナー製造プロセスで活用した結果、トナー品質管理者の工数15%低減と品質安定性の向上を確認した。

ABSTRACT

The quality control of polymerized toner, a chemical toner using polyester polymerization, is typically performed by manual adjustment. Quality operators determine the extent to which the process should be adjusted by referring to past quality values. It takes a lot of man-hours of the operators and the quality stability is insufficient. Thus, we developed an automatic quality control system, which comprises two systems: one for predicting quality and one for calculating the manipulation amount. In the quality prediction system, we selected a random forest model with variable selection from several machine learning methods. In the manipulation amount calculation system, we developed a system that automatically calculates the manipulation amount by referring to the predicted quality value. By applying the automatic quality control system to the toner manufacturing process, the quality operator's man-hours were reduced by 15% and the quality stability of the toner was improved.

-
- * CT&P本部 第2技術開発センター
2nd Technology Center, Chemical Technology & Products Division
 - ** イノベーション本部 AI応用研究センター
Applied AI Research & Development Center, Innovation/R&D Division
 - *** 生産本部
Production Division
 - **** 京都大学大学院 情報学研究所
Graduate School of Informatics, Kyoto University

第7回計測自動制御学会制御部門マルチシンポジウム, 2A2-3 (2020). 公益社団法人計測自動制御学会転載許可.

1. 背景と目的

ポリエステル重合を用いたケミカルトナーの1種である重合トナーの製造プロセスは、Fig. 1に示すように1ロットを7日かけて製造しており、毎日1ロット連続的に製造している。また、IoTを用いた製造プロセスデータ収集システムによって、原材料データや設備稼働データ、品質データなど、1ロットにつき数千項目の変数データが数百ロット収集されている。

これまで、重合トナーの品質制御には、トナー品質管理者による手動のフィードバック制御が用いられてきた。この手動制御では、制御実行時点で取得できている数ロット前の品質測定値を参照し、品質に対して感度を持つ操作変数の操作量を決定する。このような従来の品質制御では、トナー品質管理者の品質監視や操作量変更指示などに工数が多くかかるばかりでなく、トナー品質のばらつきによって規格外（品質NG）となるリスクも高く、問題が多い。そこで、トナー品質管理者の工数を削減し、品質安定性を高めるために、既存のデータ収集環境を活用してトナー品質を高精度に予測し、その予測値を活用してトナー品質を自動制御できるシステムが求められていた。

鉄鋼分野などでは、品質向上のために製造プロセスデータを用いた機械学習予測を行い、フィードフォワード制御を実用化した報告がされている¹⁾。重合トナーの品質制御においても、製造環境の変化に対応できるフィードフォワード制御が求められている。一方、事務機器の分野で機械学習を活用した事例として、複合機の故障予測²⁾やトナーの充填工程に対する予測分析³⁾などが報告されているが、トナーの製造工程への活用実績は報告されていない。そのため、本研究で我々は機械学習を用いて重合トナーの製造プロセスデータを活かした自動品質制御システムの開発を行った。これはFig. 1に示すように、製造プロセスデータを用いてトナー品質を予測する品質予測システムと、その予測値を用いて操作量を決定する操作量算出システムで構成されるフィードフォワード制御システムである。

以下、第2章では、製造プロセスデータから予測値を算出する品質予測システムについて述べる。第3章では、フィードフォワード制御を可能とする操作量算出システムを説明する。第4章では、品質予測システムと操作量算出システムで構成される品質自動制御システムをトナー製造プロセスに展開した結果を示す。第5章では、本研究のまとめを記す。

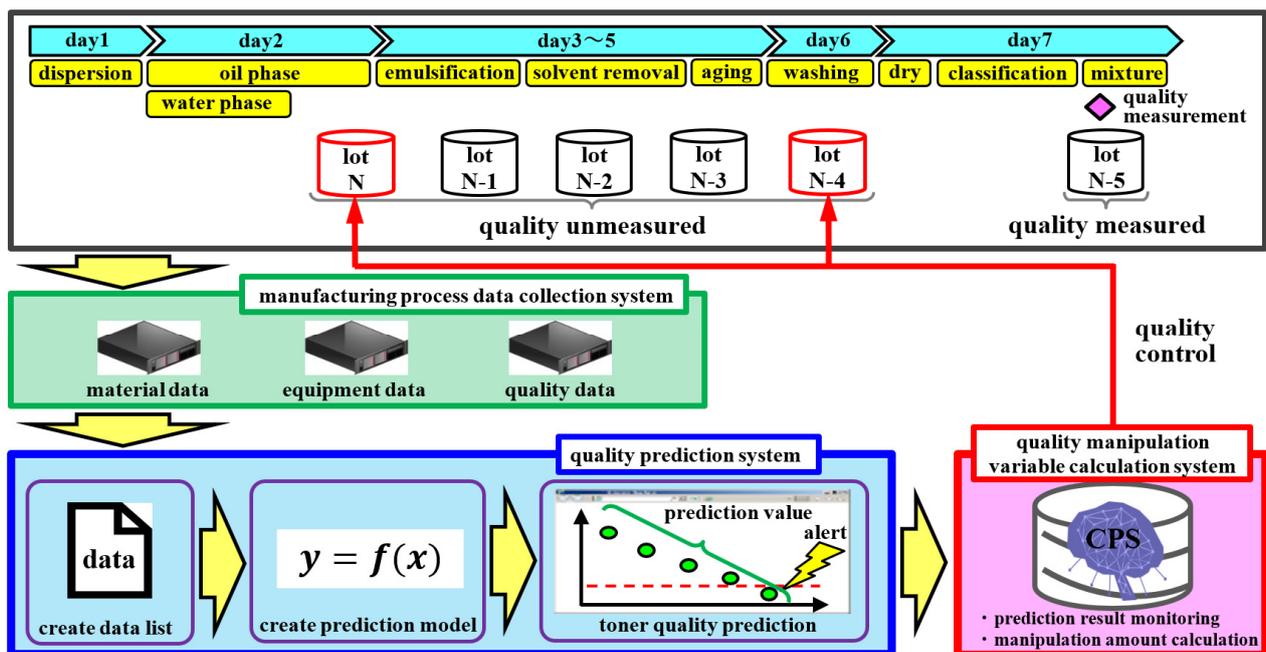


Fig. 1 Schematic diagram of automatic quality control.

2. 品質予測システム

品質予測システムには、工場から自動収集されるトナーの製造プロセスデータを用いて、高精度な予測モデルを構築することが求められる。そのため、品質と関係のある製造プロセス変数（以下「変数」と呼ぶ）を選択することと、パラメータを適切に調整することが必要である。

本章では、予測モデルの比較と、この比較に基づいて決定したRandom Forestによる予測の結果について述べる。

2-1 予測モデルの検討

予測モデルの検討では、Fig. 1に示す油相作成工程のデータを用いた。この工程では、品質制御を目的とした操作量の変更頻度が高い。

予測モデルの選択にあたっては、予測対象品質項目の全14項目のうち、測定ばらつきが少なく真値として信頼できる重要な品質項目（品質Y）を予測対象とした。また、代表的な回帰手法であるElastic Net (EN), Partial Least Squares Regression (PLSR), Support Vector Regression (SVR), Random Forest (RF), Gaussian Process Regression (GPR) の計5つのモデル構築方法⁴⁻⁸⁾を比較した。そして、選択したモデル構築方法を用いて、残り13品質項目の予測を行い、予測精度の評価を行った。なお、前処理として、変数を平均0、分散1の尺度にする正規化処理を実施した。

予測精度を評価する際には、量産運用時を想定して学習データが40ロット蓄積されてから、1ロット分増えるごとに予測モデルを更新した上で予測する、逐次モデル更新を行った。

2-2 変数選択

検討データは、ロット数が311、変数が1627項目であり、ロット数に対して変数の数が圧倒的に多いことから、過学習となりやすい。また、変数の中には品質に寄与しない項目もあり、これを取り入れて

予測モデルを構築すると、予測精度の悪化が懸念される。そのため、変数選択を実施した。

変数選択では、permutation importance⁹⁾を利用して予測モデルにおける変数重要度を算出した上で、Fig. 2のように変数重要度が高い順に変数をソートし、品質の有効桁数+1桁までの範囲で連続する3変数の移動平均値が変動しなくなるまでを重要変数として選択した。ただし、最低でも20変数は重要変数として用いた。

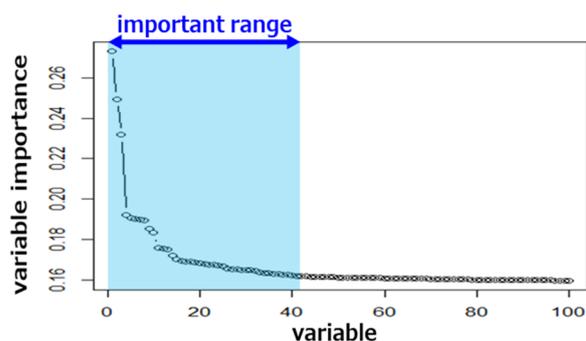


Fig. 2 Variation of variable importance.

2-3 予測精度の比較

予測精度の評価には、予測誤差RMSE (Root Mean Squared Error) の品質規格幅 Δ に対する割合を用いた。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (1)$$

$$\text{予測誤差指標} = \frac{RMSE}{\Delta} \quad (2)$$

ここで、 n はデータ総数、 \hat{y}_i は品質予測値、 y_i はFig. 1の7日目で得られる品質測定値である。予測誤差指標は、低いほうが望ましい。

また、測定値の変動に対する応答性の指標として式(3)の相関係数も用いた。

$$\text{予測応答指標} = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - Y)}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^n (\hat{y}_i - Y)^2}} \quad (3)$$

ここで、 \bar{y} は測定値の平均、 Y は予測値の平均を示す。予測応答指標は高いほうが望ましい。

311ロットのデータに対して逐次的に更新した予測モデルを用いた結果をTable 1に示す。予測誤差指

標, 予測応答指標ともに, 変数選択したRFの予測モデルが最も優れている. Fig. 3は各ロットにおける変数選択したRFの予測値 \hat{y}_i と測定値 y_i を示したグラフである. これから, 変数選択したRFは測定値を十分な精度で予測しているとともに, 許容できないほど大きな誤差が生じるロットは1つもないことが確認できた. 特に, 後者はプラントの安定運転に不可欠である. 以上のことから, 変数選択したRFを予測モデル構築手法として選定した.

Table 1 Comparison of prediction accuracy.

prediction model variable select	EN		PLSR		SVR		RF		GPR	
	all	select								
prediction error index [%]	41.0	7.5	6.5	10.9	16.4	7.4	6.7	6.2	8.8	8.4
prediction response index	0.10	0.71	0.76	0.22	0.13	0.67	0.72	0.77	0.54	0.58

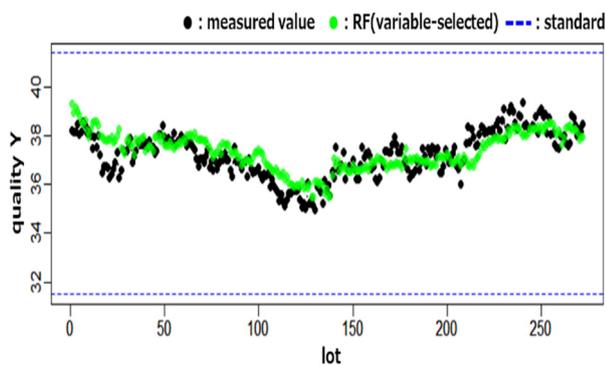


Fig. 3 Prediction results of variable-selected RF.

Table 1の結果では, PLSRは変数選択によってかえって予測誤差指標が大きくなっている. この原因を考察するため, PLSR, RFともに変数選択した予測モデルの誤差が小さいロット (PLSR: ロット59, 重要変数20項目, RF: ロット60, 重要変数29項目)を対象に, 変数重要度が高い順に変数を1つずつ増やして予測モデルを作成したときの予測誤差指標を確認した. その結果をFig. 4に示す. Fig. 4(a)を見ると, PLSRは重要変数の数が500項目近傍までは予測誤差指標が大きく変動しており, 予測誤差指標が極端に大きい重要変数の数を選択してしまうことが考えられる. それに対し, 重要変数を全て用いる800項目近傍では, 予測誤差指標の変動は小さく, 予測

誤差指標も極端に大きくなっていない. このことから, 変数選択したPLSRは全変数のPLSRよりも予測誤差指標が大きくなったと考えられる. 一方, Fig. 4(b)に見られるように, RFでは重要変数29項目を用いた場合は全ての重要変数を用いた場合と比較して予測誤差指標は小さい. このため, 変数選択した場合に予測誤差指標が小さくなったと考えられる.

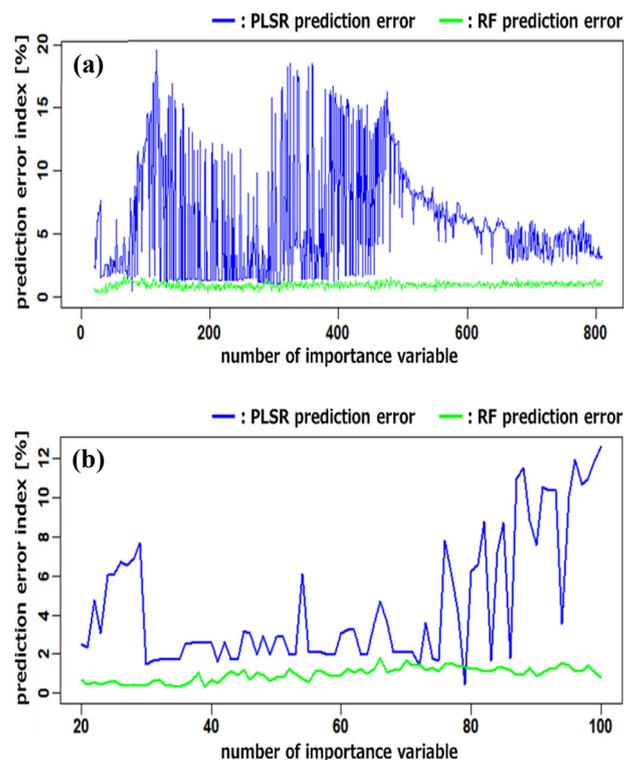


Fig. 4 Prediction error of PLSR and RF (a): up to all variables and (b): up to 100 variables.

2-4 全品質データに対する予測精度の確認

予測精度の目標としては, Fig. 5に示すように, 以下の条件を満たす範囲を設定した.

- ・ 規格の上限側と下限側において, 真値との予測誤差分布 σ_T の規格外となる確率が0.3%.
- ・ 2つの予測誤差分布 σ_T の中心の差が品質規格幅 Δ の50%.

これを満たすためには, 規格の上限側と下限側に $3\sigma_T$ 以上の余裕が求められる. すなわち,

$$0.5\Delta \geq 6\sigma_T \quad (4)$$

となる。これより、 $\sigma_T/\Delta \leq 0.5/6 = 8.3[\%]$ が予測精度の目標となる。

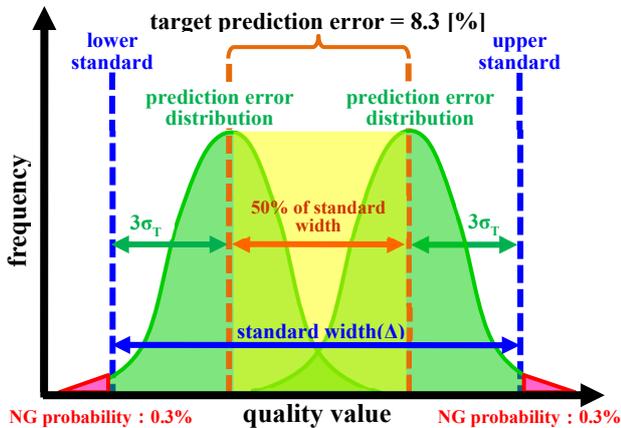


Fig. 5 Definition of interval of target prediction error.

予測誤差であるRMSEは、 σ_T と測定ばらつき σ によって以下のように分解される。

$$RMSE^2 = \sigma_T^2 + \sigma^2 \quad (5)$$

品質によってはこの測定ばらつき σ の評価が難しいため、量産データにおける予測精度の目標として、 σ_T/Δ の上限となる予測誤差指標 $RMSE/\Delta = 8.3[\%]$ を設定した。

前節で選定した変数選択したRF予測モデルが、他の13品質項目に対しても予測精度が十分であるか確認した結果をFig. 6に示す。全14品質項目が0.6%～6.2%であり、目標値8.3%以下であることから、このモデルは十分な予測精度を持っていると言える。

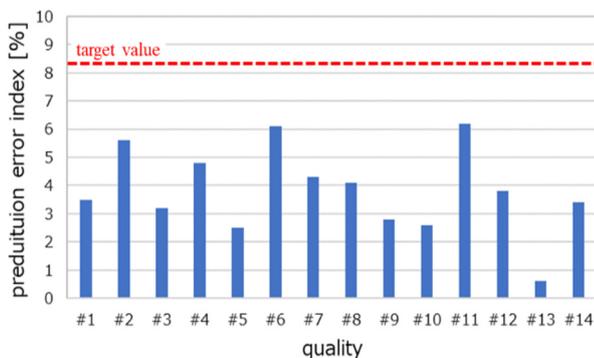


Fig. 6 Prediction accuracy for each quality.

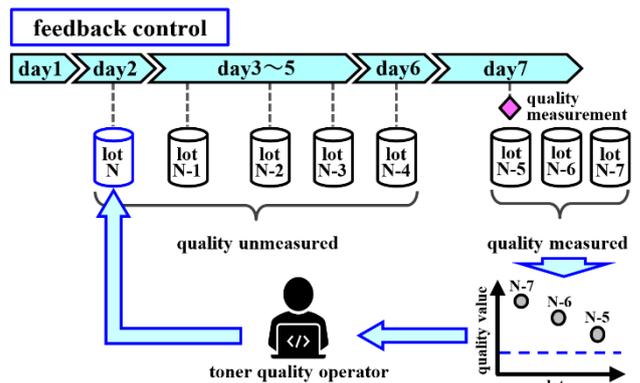
3. 操作量算出システム

操作量算出システムには、トナー品質管理者が品質制御のための操作量を決定する作業を自動化することが求められる。

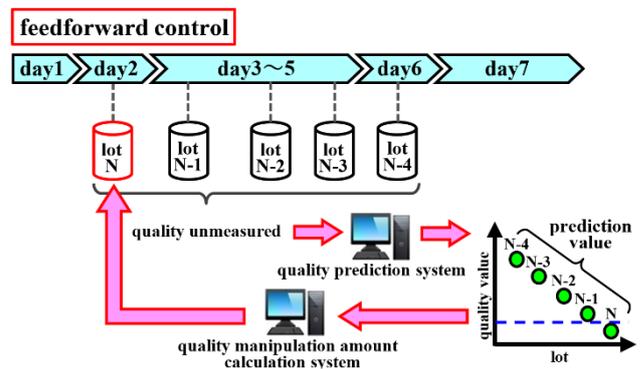
本章では、操作量算出システムとそれに用いる操作量評価関数、また、このシステムの妥当性の検証結果について述べる。

3-1 操作量算出システムの構築

制御対象工程を油相作成工程としたときの、トナー品質管理者によるフィードバック制御と機械学習を活用したフィードフォワード制御のプロセスの例をFig. 7に示す。また、その制御フローをFig. 8に示す。



Quality control for lot N based on measured quality values of lot N-5.



Quality control for lot N based on quality prediction values of lot N.

Fig. 7 Schematic diagram of feedback and feedforward control.

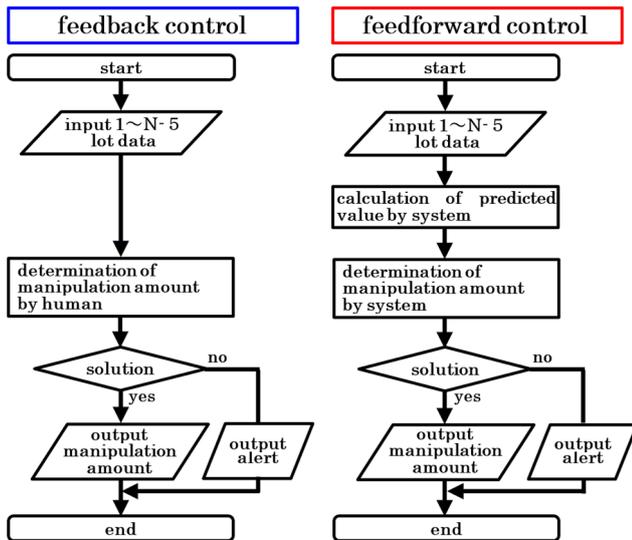


Fig. 8 Control flow of feedback and feedforward system.

フィードバック制御とフィードフォワード制御の主な違いは、フィードバック制御では第N番目のロットに注目するとき、その5ロット前のN-5ロット以前の品質測定値を参照し、トナー品質管理者が操作量を決定するのに対し、フィードフォワード制御では品質予測システムによる第Nロットの予測値を参照して操作量算出システムが操作量を決定する点である。両者とも目標品質に対して変更可能な操作量の解がある場合はその操作量を適用し、解がない場合にはアラート発報して前の工程での再調整などを行う。

3-2 操作量評価関数

トナー品質管理者によるフィードバック制御では、Table 2に示す実験で求めた感度表 S と操作変数の変更量 δx から品質変化量 δy を求める式(6)を用いて、操作量を算出している。

$$\delta y = S \cdot \delta x \quad (6)$$

ここで、 $\delta y = (\delta y_1, \dots, \delta y_p)^T$ 、 $\delta x = (\delta x_1, \dots, \delta x_q)^T$ 、

また、 p は品質項目数、 q は操作変数の数、 δx_k は操作変数の第 k 成分の操作変数の変更量、 δy_j は品質の第 j 成分の変化量、 s_{jk} は δx_k に対する δy_j の感度を示す。

Table 2 Sensitivity table for each quality.

	1	...	q
1	s_{11}	...	s_{1q}
\vdots	\vdots	\ddots	\vdots
p	s_{p1}	...	s_{pq}

操作量算出システムにおいても、この感度表に基づいた自動操作量算出を行う。制御対象工程によっては、複数の品質に対して複数の操作変数を用いて制御しなければならない場合がある。このとき、ある品質は目標値に近づいても、他の品質は目標値から離れてしまうようなトレードオフの関係となる可能性が考えられる。このため、全品質が極力目標値に近づくように操作量を求める必要がある。そこで、式(7)に与えた評価関数 f を最小化することにより操作量を決定した。

$$f(x) = \|D^{-1}(\hat{y} - \alpha + S \cdot \delta x)\|^2 \quad (7)$$

ここで、 $\hat{y} = (\hat{y}_1, \dots, \hat{y}_p)^T$ 、 $\alpha = (\alpha_1, \dots, \alpha_p)^T$ 、また、 D^{-1} は対角要素に各品質の規格幅 Δ_j を持つ対角行列である。 \hat{y}_j と α_j は、品質 y_j の予測値と目標値を示す。

式(7)は、各品質の予測値 \hat{y}_j と目標値 α_j の差と、操作量の変更による品質 y_j の変化量である $S_j \cdot \delta x_j$ が打ち消すように変更量 δx_j を探索するための評価関数である。

さらに、トナー製造において各品質が規格内に入るための品質 y_j と操作量 x_k の変更が可能な制約範囲を、制約条件(8)として設けた。

$$\left. \begin{array}{l} \text{応答量制約範囲} : y_{Cmin} \leq y_j \leq y_{Cmax} \\ \text{操作量制約範囲} : x_{Cmin} \leq x_k \leq x_{Cmax} \end{array} \right\} \quad (8)$$

3-3 操作量算出システムの妥当性確認

操作量算出システムの構築と性能評価を行った。各品質の評価のために用いる品質予測値をFig. 9に示す4水準とし、規格幅の上下限NG水準を含めるよう設定した。

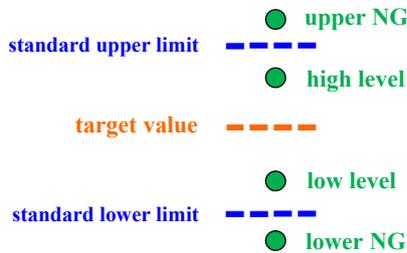


Fig. 9 Temporary target quality values for evaluating quality manipulation variable calculation system.

油相作成と洗浄の2つの工程における予測値の各水準を組み合わせた状態に対する操作量が、従来のトナー品質管理者が算出した操作量と一致するか確認した。

Table 3には油相作成工程における制御対象品質の予測値 $\hat{y}_1 \sim \hat{y}_2$ と操作変数 x_1 に対して、操作量算出システムが算出した最適操作量と、トナー品質管理者が算出した操作量の一部を示した。他の水準を含め、両者は全て一致することが確認できた。

Table 3 Evaluation of operational value calculation system for oil-phase creation process.

prediction value		manipulation amount calculation system		feedback control by operator	
quality 1 [\hat{y}_1]	quality 2 [\hat{y}_2]	solution	manipulation amount [x_1] (-150 $\leq x_1 \leq 150$)	solution	manipulation amount [x_1] (-150 $\leq x_1 \leq 150$)
level 1	level 1	no	-	no	-
level 2	level 1	yes	-150	yes	-150
level 3	level 1	yes	-150	yes	-150
level 4	level 1	yes	-150	yes	-150
level 1	level 2	yes	100	yes	100
⋮	⋮	⋮	⋮	⋮	⋮

Table 4には、洗浄工程における制御対象品質の予測値 $y_3 \sim y_5$ と操作変数 $x_2 \sim x_3$ に対して、操作量算出システムが算出した最適操作量と、トナー品質管理

者が算出した操作量を示した。洗浄工程も両者が完全に一致することを確認した。

Table 4 Evaluation of operational value calculation system for cleaning process.

prediction value			manipulation amount calculation system		feedback control by operator		
quality 3 [\hat{y}_3]	quality 4 [\hat{y}_4]	quality 5 [\hat{y}_5]	solution	manipulation amount [x_2] and [x_3] (24 $\leq x_2 \leq 45$) (0 $\leq x_3 \leq 30$)	solution	manipulation amount [x_2] and [x_3] (24 $\leq x_2 \leq 45$) (0 $\leq x_3 \leq 30$)	
level 1	level 1	level 1	yes	38 10	yes	38 10	
level 2	level 1	level 1	yes	40 10	yes	40 10	
level 3	level 1	level 1	yes	40 10	yes	40 10	
level 4	level 1	level 1	no	- -	no	- -	
level 1	level 2	level 1	yes	25 0	yes	25 0	
⋮	⋮	⋮	⋮	⋮	⋮	⋮	

これらの結果から、従来のトナー品質管理者と同等の操作量を算出できる操作量算出システムが構築できた。

4. 品質自動制御システムの評価

第2章、第3章に述べた予測モデルと操作量算出システムを組み合わせた品質自動制御システムを構築し、重合トナーの量産における評価を実施した。品質自動制御システム導入前後41ロットを評価対象とし、品質達成能力を表す式(9)の工程能力指数 (C_{pk}) を評価指標とした。

$$K = \frac{|(USL + LSL)/2 - \mu|}{(USL - LSL)/2}, \quad C_p = \frac{USL - LSL}{6\sigma_c}.$$

$$C_{pk} = (1 - K)C_p \quad (9)$$

ここで、 USL は上限規格値、 LSL は下限規格値、 μ は工程測定値の平均、 σ_c は工程測定値の標準偏差を示す。 C_{pk} は不良率の最大値を示しており、1であれば最大不良率が0.27%、1.33であれば最大不良率が0.006%となる。一般的に、 C_{pk} が1未満であれば工程能力はなく、1以上1.33未満であれば工程能力はいま一步、1.33以上であれば工程能力はあるとされている¹⁰⁾。

一例として、油相工程の制御対象品質である品質1と品質2の品質自動制御システム導入前後における工程能力指数の結果をTable 5に示す。精度の高い予測と、その予測値に基づいた適切な品質制御を毎ロット自動で実施した結果、品質1、品質2ともに従来の手動フィードバック制御以上の工程能力指数となり、トナー品質管理者を介さずに従来以上の品質達成能力を持ったシステムが構築できた。

Table 5 Improvement in process capability index before and after introduction of automatic quality control system.

	process capability index (C_{pk})	
	quality 1	quality 2
41 lots immediately before	1.07	1.41
41 lots immediately after	1.28	1.49

この品質自動制御システムを重合トナーの製造プロセスに展開した結果、トナー品質管理者の工数15%の低減と、工程能力指数 (C_{pk}) の向上による品質安定化に寄与することができた。

5. 結論

今回対象とした重合トナーに対し、機械学習予測技術を用いた品質予測システムと操作量算出システムを組み合わせて、フィードフォワード制御が可能な品質自動制御システムを構築し、それを評価した。その結果、従来のフィードバック制御時と比べて、トナー品質管理者の工数15%低減と品質安定性の向上を確認した。

今後は、品質自動制御システムを水平展開し、更なる工数低減と品質安定性の向上を目指すとともに、メンテナンスも容易なシステムへと改善していきたい。特に、後者については感度表を自動更新する技術の開発に取り組みたい。

参考文献

- 1) 茂森弘靖ほか: 制約付き局所回帰モデルを用いた鋼材の平面形状制御, 計測自動制御学会論文集, Vol. 46, No. 8, pp. 472–479 (2010).
- 2) 木村俊一ほか: 人とAIの協働によるデータエントリー業務改革, 富士ゼロックステクニカルレポート, No. 27, pp. 16–22 (2018).
- 3) Ricoh UK and Ireland: Predictive Analytics for Manufacturing, Ricoh Products Ltd. Case Study, <https://www.youtube.com/watch?v=nZO388RNhQg> (accessed 2019-11-08).
- 4) H. Zou, T. Hastie: Regularization and Variable Selection Via the Elastic Net, *Journal of the Royal Statistical Society*, Vol. 67, No. 2, pp. 301–320 (2005).
- 5) S. Wold, M. Sjöström, L. Eriksson: PLS-Regression: A Basic Tool of Chemometric, *Chemometrics and Intelligent Laboratory Systems*, Vol. 58, pp. 109–130 (2001).
- 6) A.J. Smola, B. Schölkopf: A Tutorial on Support Vector Regression, *Statistics and Computing*, Vol. 14, pp. 199–222 (2004).
- 7) L. Breiman: Random Forests, *Machine Learning*, Vol. 45, pp. 5–32 (2001).
- 8) C. E. Rasmussen, C. K. I. Williams: Gaussian Processes for Machine Learning, The MIT Press, pp. 7–32 (2006).
- 9) A. Althmann et al.: Permutation Importance: A Corrected Feature Importance Measure, *Bioinformatics*, Vol. 26, Issue 10, pp. 1340–1347 (2010).
- 10) 永田靖, 棟近雅彦: 工程能力指数 実践方法とその理論, 一般財団法人 日本規格協会, pp. 17–35 (2011).