

近赤外分光法による医薬品品質管理技術と  
その継続的改善手法に関する研究

Research on pharmaceutical quality control technique by  
near-infrared spectroscopy and its continuous improvement

博士（薬学）学位論文

東京薬科大学大学院薬学研究科

2016 年

研究分野 薬剤学

指導教授 瀬田 康生

学位申請者 宮野 拓也

本研究は、以下の原著論文に基づいたものである。

- 1) Miyano, T., Kano, M., Tanabe, H., Nakagawa, H., Watanabe, T., Minami, H., 2014. Spectral fluctuation dividing for efficient wavenumber selection: Application to estimation of water and drug content in granules using near infrared spectroscopy. *Int. J. Pharm.* 475, 504–513. (第一章)
- 2) Miyano, T., Fujiwara, K., Kano, M., Tanabe, H., Nakagawa, H., Watanabe, T., Minami, H., 2015. Efficient wavenumber selection based on spectral fluctuation dividing and correlation-based clustering for calibration modeling. *Chemom. Intell. Lab. Syst.* 148, 85–94. (第二章)
- 3) Miyano, T., Nakagawa, H., Watanabe, T., Minami, H., Sugiyama, H., 2015. Operationalizing maintenance of calibration models based on near-infrared spectroscopy by knowledge integration. *J. Pharm. Innov.* 10, 287–301. (第三章)

## 目次

緒言 .....	1
略号表 .....	8
第一章 スペクトル変動分割を用いた効率的な波数選択法の開発 .....	12
1.1 序論 .....	12
1.2 実験の部 .....	13
1.2.1 材料 .....	13
1.2.2 NIRS 測定 .....	13
1.2.3 参照測定 .....	14
1.2.4 PLS .....	14
1.2.5 キャリブレーションセット及びバリデーションセット並びに推定精度の指標 .....	16
1.2.6 スペクトル前処理 .....	16
1.2.7 計算時間 .....	17
1.3 波数選択法 .....	17
1.3.1 PLS-beta .....	18
1.3.2 VIP .....	18
1.3.3 iPLS .....	19
1.3.4 SFD-PLS (提案法) .....	20
1.3.5 MASFD-PLS (提案法) .....	22
1.3.6 SFDA-SLR (提案法) .....	24
1.4 結果及び考察 .....	25
1.4.1 スペクトル前処理 .....	25
1.4.2 スペクトル変動分割 .....	25
1.4.3 波数選択及び検量モデル構築 .....	27
1.4.3.1 水分含量推定 .....	27
1.4.3.2 薬物含量推定 .....	29
1.4.4 計算時間 .....	32
1.5 小括 .....	33
第二章 相関に基づくクラスタリングを用いた効率的な波数選択法の開発 .....	35
2.1 序論 .....	35
2.2 実験の部 .....	36
2.2.1 材料 .....	36
2.2.2 NIRS 測定 .....	36
2.2.3 参照測定 .....	36
2.2.4 PLS .....	36
2.2.5 キャリブレーションセット及びバリデーションセット並びに推定精度の指標 .....	36

2.2.6	スペクトル前処理	37
2.2.7	計算時間	37
2.3	波数選択法	37
2.3.1	iPLS	38
2.3.2	SCMWPLS	39
2.3.3	SFD-PLS	39
2.3.4	<i>k</i> -means-PLS	39
2.3.5	NCSC-PLS	40
2.3.6	Interval- <i>k</i> -means-PLS	41
2.3.7	Interval-NCSC-PLS	41
2.3.8	SFD- <i>k</i> -means-PLS	42
2.3.9	SFD-NCSC-PLS (提案法)	42
2.4	結果及び考察	44
2.4.1	波数選択及び検量モデル構築	44
2.4.3.1	水分含量推定	44
2.4.3.2	薬物含量推定	48
2.4.2	計算時間	52
2.5	小括	54
<b>第三章 検量モデルメンテナンスの知識統合に基づく運用化</b>		<b>57</b>
3.1	序論	57
3.2	実験の部	58
3.2.1	業務プロセスモデル化技術: IDEF0	58
3.2.2	役割分担表: RACI	59
3.2.3	知識統合の効果を評価する指標	60
3.3	結果及び考察	61
3.3.1	検量モデルメンテナンスの枠組み	61
3.3.1.1	メンテナンスの必要性判断	62
3.3.1.2	メンテナンス実施	64
3.3.2	業務プロセスモデル化に基づく知識統合	66
3.3.2.1	知識統合前の検量モデルメンテナンス	66
3.3.2.2	業務プロセスモデルの作成	67
3.3.2.3	IDEF0 モデルの解釈	69
3.3.2.4	RACI 表の解釈	72
3.3.3	検量モデルメンテナンスの運用化	73
3.3.3.1	検量モデルメンテナンス SOP の更新	73
3.3.3.2	作業効率の改善	75
3.3.3.3	検量モデルメンテナンス運用の検証	79

3.4 小括.....	80
総括 .....	81
謝辞 .....	84
引用文献 .....	85

## 緒言

医薬品の品質をより高度に保証するためのパラダイムとして Quality by Design (QbD) 及びプロセス解析技術 (process analytical technology : PAT) が注目されており、これら QbD 及び PAT に関するガイダンスが米国食品医薬品局 (FDA)<sup>1)</sup>、欧州医薬品庁 (EMA)<sup>2), 3), 4)</sup>、及び日米 EU 医薬品規制調和国際会議 (ICH)<sup>5), 6), 7)</sup>により発出されている。QbD は、製造プロセスと製品品質との関連を明らかにし、設計及び開発段階で品質を造り込むことを指向した枠組みである。QbD によって得られた知識に基づいて品質リスクを管理することで、適切な品質管理戦略を構築することができる。PAT は、物質特性を迅速に測定する技術であり、QbD を実現するために製造プロセスを設計、解析、及び制御する手段として利用される。製薬産業は PAT を活用することで、製品品質及び生産効率を向上させることに成功している<sup>8), 9)</sup>。

PAT として、近赤外分光法 (near-infrared spectroscopy: NIRS) が一般的に用いられている。近赤外 (near-infrared: NIR) 領域は、可視光領域と赤外領域には含まれた  $12500\text{ cm}^{-1}$  から  $4000\text{ cm}^{-1}$  の波数領域の光である。NIRS では、試験サンプルを溶解する等の前処理が不要及び迅速かつ非破壊の測定が可能であるため、製造工程中での迅速ないしリアルタイムの品質管理に適している。また、その測定対象は錠剤、顆粒、粉体、液体、及び懸濁液等と幅広く利用されている。さらに、NIR スペクトルは様々な化学的及び物理的特性を反映するため、様々な品質特性を管理するために NIRS が用いられている<sup>10), 11)</sup>。固形製剤の製造にあたり NIRS を用いて管理する品質特性の例として、造粒工程における顆粒中の水分含量<sup>12), 13)</sup>、混合工程における混合均一性<sup>14), 15)</sup>、コーティング工程におけるフィルム膜厚<sup>16), 17)</sup>、及びサンプリングした顆粒または錠剤中の薬物含量<sup>18), 19)</sup>が挙げられる。ただし、NIR スペクトルはこうした様々な化学的及び物理的特性に関する情報が入り混じったものである。また、NIR スペクトルの挙動は複雑である。複雑な挙動を示す理由として、一つの官能基の振動モードに対して一般に複数の NIR 吸収帯が存在すること、各 NIR 吸収帯に応じた光の吸収は微弱であり、各吸収ピークは広範な形状となること、混合物においてはそれぞれの化合物に由来する広範な吸収ピークが重なり合うこと、隣接する官能基や測定対象の温度に応じて NIR 吸収帯の位置が変動すること、及び測定対象の密度や空隙といった物理的性質がスペクトル強度に影響を与えること等が挙げられる。また、Fig. 1 に示すように、NIR スペクトルにおける単一の波数のみで、目的とする品質特性を推定することは困難である。そこで、NIR スペクトルから目的とする品質特性に関する情報を得るためには、NIR スペクトルにおける複数の波数点と品質特性との関係を表す検量モデルが必要となる。

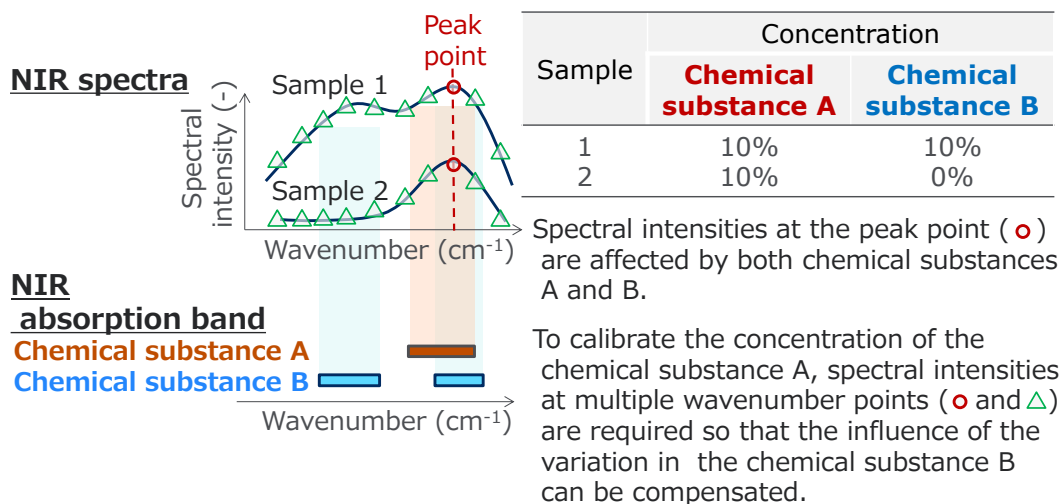


Fig. 1 An illustrative example of overlapping of NIR absorption bands.

検量モデルを構築するために用いる入力変数及び出力変数のデータセットをキャリブレーションセットと呼ぶ。ここで入力変数とは NIR スペクトルのスペクトル強度（ただし、後述するスペクトル前処理を施した後のスペクトル強度を用いる）であり、出力変数とは目的とする品質特性（第一章及び第二章の事例研究では、水分含量または薬物（化合物 X）含量）である。これらの入力変数及び出力変数に対して回帰分析を実施することで検量モデルが得られる。なお、HPLC 等の参照法で測定することで得た品質特性を出力変数の参照値、検量モデルを用いた回帰分析によって得られた品質特性を出力変数の推定値と呼ぶ。

Fig. 2 に示すように、検量モデルを構築する手順は以下の通りである。

1) キャリブレーションセット用のサンプルを選択する

目的とする品質特性を意図的に変動させたサンプルを用いる（例えば、薬物含量を 70% から 130% の範囲で変動させる）。また、目的以外の品質特性（例えば、薬物含量を推定する検量モデルにおける水分含量）を意図的に変動させたサンプルを採用することで、その目的以外の品質特性に関する情報を予め検量モデルに組み込み、その変動が検量モデルの推定結果に与える影響を低減させることができる。なお、NIR スペクトルに影響を与えるあらゆる物質特性を把握することは困難であるため、それぞれのサンプルには意図しない物質特性の変動が必然的に含まれていることになる。

2) キャリブレーションセット用のサンプルを NIRS 及びその参照法で測定する

NIRS は非破壊測定であるため、同一のサンプルを NIRS と参照法の両方で測定することができる。NIRS で測定することで得た NIR スペクトルに対し、後述のスペクトル前処理及び波数選択を実施した後のスペクトル強度を入力変数として用いる。参照法で測定することで得た品質特性の値を出力変数の参照値として用いる。

### 3) NIR スペクトルにスペクトル前処理を適用する

NIR スペクトルにスペクトル前処理を施すことで、物理的な外乱因子の影響を低減し、吸収ピークを顕在化させることができる。スペクトル前処理の詳細は「1.2.6 スペクトル前処理」を参照のこと。

### 4) 使用する波数を選択する

NIR スペクトルには目的とする品質特性と関係がある波数だけではなく、目的とする品質特性と無関係な波数が含まれる。この無関係な波数を検量モデルに採用すると、その推定精度が低下する場合と向上する場合の両方の可能性がある。推定精度が低下するのは、この無関係な波数を検量モデルに採用することで目的外の物質特性の変動が外乱情報として検量モデルの推定結果に影響する場合である。推定精度が向上するのは、この無関係な波数を採用することで目的とする品質特性と無関係な特性の変動が推定精度に与える影響を適切に補正できる場合である (Fig. 1 を参照)。従って、推定精度の高い検量モデルを構築する上で、目的に応じた適切な波数を選択することは必須である。スペクトル前処理後の NIR スペクトルにつき、選択した波数のみを次の回帰分析に用いる。

### 5) 回帰分析を実施する

NIR スペクトルは  $12500\text{ cm}^{-1}$  から  $4000\text{ cm}^{-1}$  の範囲における数千の波数点のスペクトル強度で構成される。この膨大な数のスペクトル強度を入力変数として検量モデルを構築する際、重回帰分析法を用いると入力変数間の相関性が原因となって推定精度が著しく低下する多重共線性の問題が発生する。この多重共線性に対応するために、部分最小二乗法 (partial least squares: PLS) が多変量回帰分析法として広く採用されている。



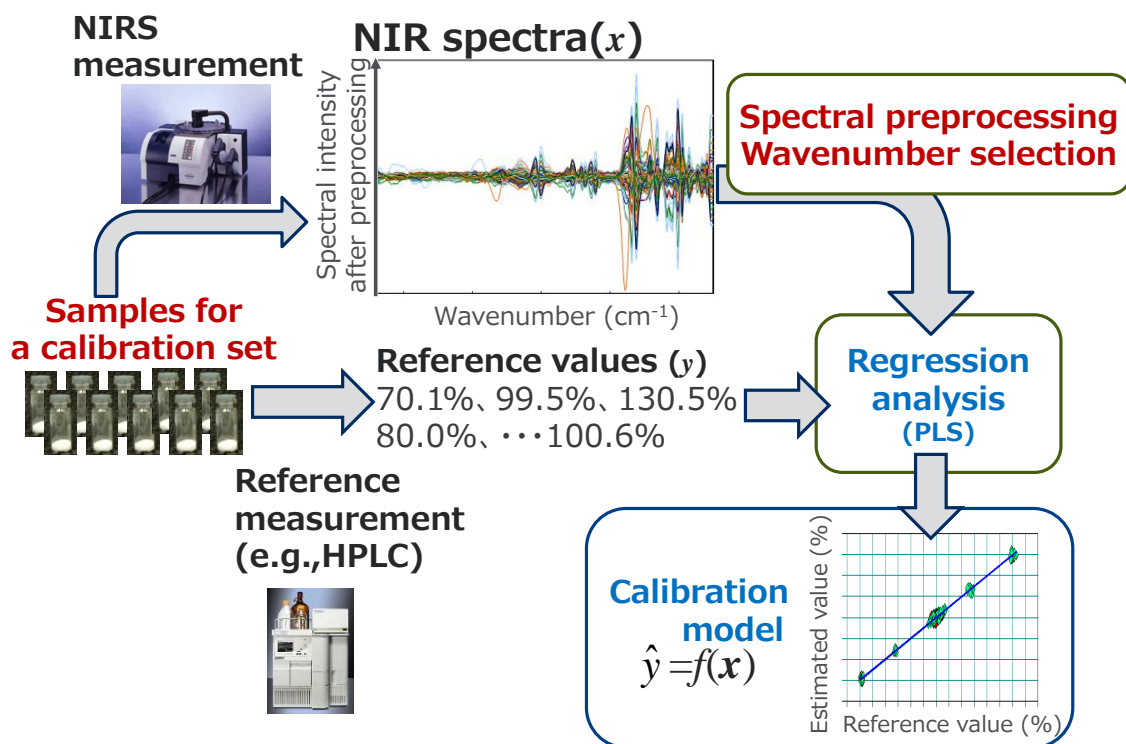


Fig. 2 Workflow to develop a calibration model.

検量モデルの最も重要な性能指標は推定精度である。検量モデルの推定精度を検証するために用いる入力変数及び出力変数のデータセットをバリデーショナルセットと呼ぶ。バリデーショナルセットは、キャリブレーションセットとは独立したデータセットとする。推定精度を検証する際は、検量モデルの構築と同様の手順で、バリデーショナルセット用のサンプルを NIRS 及びその参照法で測定する。次に、得られた NIR スペクトルに対し、検量モデルで採用したスペクトル前処理を適用する。そのスペクトル前処理後の NIR スペクトルにつき、検量モデル構築で選択した波数のみを入力変数として用いる。その入力変数を検量モデルに当てはめ、出力変数の推定値を得る。出力変数の推定値及び参照値を比較することで、推定精度を検証する。推定精度の評価指標は後述の「1.2.5 キャリブレーションセット及びバリデーショナルセット並びに推定精度の指標」を参照のこと。

推定精度は採用する多変量回帰分析法、キャリブレーションセット用のサンプル、NIR スペクトルの前処理、及び波数に依存する。それ故、推定精度を向上させるために、多変量回帰分析法<sup>20)</sup>、キャリブレーションセット用のサンプル選択<sup>21), 22), 23)</sup>、NIR スペクトルの前処理<sup>24), 25), 26)</sup>、そして波数選択<sup>27), 28)</sup>に関する各種の研究が精力的に実施されている。

高い推定精度を有する検量モデルを開発する上で、波数選択は特に重要な役割を担う。これは品質特性と無関係な波数、つまり品質特性の変動に対して相対的に外乱変動の影響が大きい波数を選択すると、外乱変動の影響を受けて推定誤差が増大する

場合がある為である。波数選択を実施する方策は以下の三つが考えられる。①波数の全組合せを評価する、②分光学的知識に基づいて波数を選択する、及び③統計的な手法に基づいて波数を選択する。①波数の全組合せを評価する場合、その組合せ数は膨大であるため、非常に長い計算時間を要する（第一章及び第二章で用いた条件では、NIR スペクトルは 2202 点の波数で構成されるため、 $2^{2202}$  通りの組合せとなる。この全ての組合せについて検量モデルを構築するとおよそ  $10^{658}$  年かかる）。従って、①は非現実的である。②分光学的知識に基づいて波数を選択することは可能である。しかし、NIR スペクトルにおいては複数の官能基に基づく吸収帯が重なり合っており、さらにスペクトル強度が顆粒の粒子径等の物理的特性の影響を受けるといった性質があるため、その品質特性の変動を適切に反映する波数を分光学的知識のみに基づいて特定することは困難である。従って、試行錯誤の評価が必要になる為、作業工数が大きい及び適切な波数が選択できるか否かは個人の能力や運に依存するといった課題がある。そこで、③統計的な手法に基づいて波数を選択することが提案されてきた。統計的な波数選択法の目的は、高い推定精度を得る波数を短い計算時間で選択することである。計算時間は NIR スペクトルのデータ量（サンプル数及び波数点の数）、計算機の性能、及び波数選択アルゴリズムに依存する。しかし、従来の統計的な波数選択法におけるアルゴリズムは計算負荷が非常に大きいため、NIR スペクトルのデータ量によっては高性能な計算機を用いたとしても膨大な計算時間を要することが課題として認識されている。また、推定精度を改善する観点で、波数選択アルゴリズムを改善できる余地がある。

検量モデルを開発した時点において検量モデルが所望の推定精度を有していたとしても、医薬品のライフサイクルを通じてその推定精度が恒常的に維持されるとは限らない。これは NIR スペクトルが推定対象（e.g., 薬物）以外の化学物質（e.g., 水分）に関する情報及び物理的な情報（e.g., 測定物質の密度）を反映するためである。従って、開発段階では NIR スペクトルに影響を与える変動因子を予め検証し、推定対象以外の変動因子の影響を受けにくい頑健な検量モデルを構築する。しかし、医薬品生産において原料や製造プロセスの特性が想定する範囲を超えて変動する可能性がある。例えば、製剤の原料である添加剤の供給業者を変更すると、その添加剤の物理的特性が変動し、それを用いて製造される製剤の物理的特性に影響を与え得る。このような外乱変動は製剤の NIR スペクトルに影響を与え、それによって NIRS に基づく推定結果に影響を与える可能性がある<sup>29)</sup>。実際、製造プロセスの特性が変動することによって検量モデルの推定精度が低下することが検量モデルを用いた品質管理における主要課題として認識されている<sup>30)</sup>。こうした課題に対応するために、検量モデルをメンテナンスすること、つまり継続的に推定精度を検証し、必要に応じて検量モデルを更新することで外乱変動に対する頑健性を逐次向上させるという方策が講じられている。この検量モデルメンテナンスを商用生産工場で適切に運用することが、検量モデルの推定精度を継続的に保証する上で重要である。

本論文では、NIRS に基づく検量モデルの推定精度を効率的に向上させるための波数選択法を提案する。また、検量モデルメンテナンスを知識統合に基づいて効率的かつ適切に運用化する手法を提案する。

第一章では、化学物質の含量が変動するとその NIR 吸収帯に対応する特定の吸収ピーク（スペクトル領域におけるスペクトル強度）に影響を与えることに着目し、新規のスペクトル分割法に基づく波数選択法を提案する。ここでスペクトル領域とは、ある幅を有する波数領域を意味する。従来の波数選択法である interval PLS (iPLS) はスペクトル全体を等幅のスペクトル領域に分割し、そのスペクトル領域単位で波数を選択する。これは吸収ピークに対応するスペクトル領域を考慮した手法であるが、等幅のスペクトル領域では推定精度を向上させる上で有用な波数を柔軟に選択することができない。そこで、吸収ピークに対応するスペクトル領域をより正確に検出する手法を提案する。吸収ピークの境界ではスペクトル強度の変動が極小となると考えられる。提案するスペクトル分割法であるスペクトル変動分割 (spectral fluctuation dividing: SFD) は、キャリブレーションセットにおける各波数点のスペクトル強度の標準偏差を算出し、スペクトル変動特性とする。そして、そのスペクトル変動特性の極小値でスペクトル全体を複数のスペクトル領域に分割する。SFD に基づく波数選択法は、SFD で作製したスペクトル領域単位で波数を入力変数として使用するか否かを選択する。これによって、推定精度を向上させる上で有用な波数をより適切に選択できることが期待される。提案する波数選択法を用いて顆粒中の水分含量及び薬物（化合物 X）含量を推定する 2 種類の検量モデルを構築し、推定精度及び計算時間の観点で従来の波数選択法と比較評価した。

第二章では、化学物質が一つの官能基の振動モードに対して一般に複数の NIR 吸収帯を有しており、複数のスペクトル領域におけるスペクトル強度が類似した挙動を示すという NIR スペクトルの特徴に着目し、スペクトル強度の相関性を利用した新規のクラスタリングに基づく波数選択法を提案する。第一章で提案した SFD-PLS は各吸収ピークに対応するスペクトル領域を作製することを意図したものである。しかし、そのスペクトル領域の数は膨大であり、より効率的に高い推定精度を得る観点で改善の余地がある。第二章で提案するクラスタリング法である SFD-nearest correlation spectral clustering (SFD-NCSC) は、SFD で作製したスペクトル領域におけるスペクトル強度の相関に基づいてスペクトル領域をクラスタリングし、スペクトル領域グループを作製する。SFD-NCSC に基づく波数選択法は、SFD-NCSC で作製したスペクトル領域グループ単位で波数を選択する。これによって、一つの変動因子に由来する複数の吸収ピークに対応するスペクトル領域グループ単位で波数を選択できることが期待される。提案する波数選択法を用いて顆粒中の水分含量及び薬物（化合物 X）含量を推定する検量モデルを構築し、推定精度及び計算時間の観点で従来の波数選択法と比較評価した。

第三章では、商用生産工場における検量モデルメンテナンスを知識統合に基づい

て運用化することを提案する。検量モデルメンテナンスの重要性及び特定の評価手法が概念としては報告されているものの<sup>30), 31), 32)</sup>、その具体的な枠組み及び運用化に向けた方法論はこれまで報告されていない。そこで、まず各種のガイドライン<sup>31), 32)</sup>及び米国薬局方 (USP)<sup>33)</sup>を参考に検量モデルメンテナンスの具体的な枠組みを新規に構築した。この枠組みを商用生産工場にて運用する場合、製造管理、品質管理、品質保証、及び技術開発等の複数の部署が連携して作業を実施する。また、適正製造規範 (GMP) を遵守する必要がある。GMP 管理下では、詳細な作業手順及び役割を明確化し、文書化することが求められる。しかし、複数の部署が関与する新規の枠組みに対して、あらゆる事態に対応する詳細な作業手順及び役割を予め明確化することは困難である。そこで、検量モデルメンテナンスを適切かつ効率的に運用化する方法論として、各部署が有する知識を統合することを提案する。ここで知識を統合するとは、個々の担当者が有する知識を関連付け、文章や図表を用いた有形の知識にまとめることによって、あらゆる担当者が個々の担当者が有する知識を活用できる状態にすることである。この知識統合によって、複数の部署が関与する新規の枠組みに対して、あらゆる事態に対応する詳細な作業手順及び役割を予め明確化できることが期待された。知識を統合する手段として、業務プロセスモデル化技術である 0 型統合化定義方法論 (the type zero method of the integrated definition language: IDEF0) 及び役割分担表 RACI を用いた。また、知識統合が検量モデルメンテナンスを運用する効率に与える効果を評価するとともに、商用生産工場において検量モデルメンテナンスが適切に運用化されることを検証した。

## 略号表

本文では以下の略号を使用する。

$0$	The zero matrix, 零行列
$A_{ni}$	Interval area at the $i$ th spectral interval in the $n$ th sample, $n$ 番目のサンプルにおける $i$ 番目のスペクトル領域の領域面積
$b$	Weighting vector of $y$ , $y$ の重み付けベクトル
$b_k$	The $k$ th element in weighting vector of $y$ , $y$ の重み付けベクトルの $k$ 番目の値
$\beta_{\text{PLS}}$	Regression coefficient vector, 回帰係数ベクトル
$C_{k,l}$	Correlation coefficient between $x_{mk}$ and $x_{ml}$ , $x_{mk}$ と $x_{ml}$ との間の相関係数
CSMWPLS	Changeable size moving window partial least squares
$\Delta W_{\text{mov}}$	Moving window interval, 移動枠の刻み幅
$E$	Residual matrix of $X$ , $X$ の残差行列
EMA	European Medicines Agency, 欧州医薬品庁
$f$	Residual vector of $y$ , $y$ の残差ベクトル
FDA	Food and Drug Administration, 食品医薬品局
GA-PLS	Genetic algorithm-partial least squares
$G_{\text{gen}}$	The number of generated wavenumber groups, 作製する波数グループの数
$G_{\text{max}}$	The maximal value of $G_{\text{gen}}$ , $G_{\text{gen}}$ の最大値
$G_{\text{min}}$	The minimal value of $G_{\text{gen}}$ , $G_{\text{gen}}$ の最小値
GMP	Good manufacturing practice, 適正製造規範
$G_{\text{sel}}$	The number of selected wavenumber groups, 選択する波数グループの数
$G_t$	The tentative value of $G_{\text{gen}}$ , $G_{\text{gen}}$ の暫定値
HPLC	High performance liquid chromatography, 高速液体クロマトグラフィー
ICH	International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use, 日米EU医薬品規制調和国際会議
$i$	The position of the spectral interval, スペクトル領域の位置
IDEF0	The type zero method of the integrated definition language, 0型統合化定義方法論
$I_{\text{div}}$	The number of the divided spectral intervals, 分割されたスペクトル領域の数
$I_{\text{max}}$	The maximal value of $I_{\text{div}}$ , $I_{\text{div}}$ の最大値
$I_{\text{min}}$	The minimal value of $I_{\text{div}}$ , $I_{\text{div}}$ の最小値
iPLS	Interval PLS
$I_{\text{sel}}$	The number of the selected spectral intervals, 選択されたスペクトル領

	域の数
$K$	The number of adopted latent variables, 採用した潜在変数の数
LOD	Loss on drying, 乾燥減量
LOOCV	Leave-one-out cross validation, 1つ抜きクロスバリデーション
$M$	The number of wavenumbers, 波数の数
$M_{\text{all}}$	The number of wavenumbers in a whole spectrum, スペクトル全体における波数の数
MASFD	Moving averaged spectral fluctuation dividing, 移動平均スペクトル変動分割
MASFD-PLS	Moving averaged spectral fluctuation dividing-partial least squares, 移動平均スペクトル変動分割-部分最小二乗法
$M_i$	The number of wavenumbers in the $i$ th spectral interval, $i$ 番目のスペクトル領域における波数の数
MWPLS	Moving window partial least squares
MSPC	Multivariable statistical process control, 多変量統計工程解析
$N$	The number of samples, サンプルの数
NCSC	Nearest correlation spectral clustering
NIPALS	Nonlinear iterative partial least squares, 非線形反復部分最小二乗法
NIR	Near-infrared, 近赤外
NIRS	Near-infrared spectroscopy, 近赤外分光法
$P$	Loading matrix of $X$ , $X$ のローディング行列
$P$	The number of initialization in $k$ -means method, $k$ -means 法における初期化回数
PAT	Process analytical technology, プロセス解析技術
Ph. Eur.	European Pharmacopoeia, 欧州薬局方
$p_k$	The $k$ th loading vector of $X$ , $X$ の $k$ 番目のローディングベクトル
$P_{\text{max}}$	Predetermined maximal value of $P$ , 予め設定した $P$ の最大値
PLS	Partial least squares, 部分最小二乗法
$Q$	The number of assignment in $k$ -means method, $k$ -means 法における割当て回数
QbD	Quality by design, クオリティバイデザイン
$Q_{\text{max}}$	Predetermined maximal value of $Q$ , 予め設定した $Q$ の最大値
$R^2$	Coefficient of determination, 決定係数
$R_{\text{thres}}$	Threshold of correlation coefficient, 相関係数の閾値
$S$	Affinity matrix, 類似度行列
SCMWPLS	Searching combination moving window partial least squares
SEC	Standard error of calibration, キャリブレーションの標準誤差

SECV	Standard error of cross validation, クロスバリデーションの標準誤差
SEP	Standard error of prediction, 予測標準誤差
SFD	Spectral fluctuation dividing, スペクトル変動分割
SFDA	Spectral fluctuation dividing area, スペクトル変動分割面積
SFDA–SLR	SFD area–simple linear regression, スペクトル変動分割面積–単回帰分析 析法
$S(k, l)$	The $(k, l)$ element of $S$ , $S$ の $k$ 行 $l$ 列目の要素
SLR	Simple linear regression, 単回帰分析法
$s_m$	Standard deviation of $\mathbf{x}_m$ , $\mathbf{x}_m$ の標準偏差
SNV	Standard normal variate, 標準正規変量
SOP	Standard operating procedure, 標準操作手順書
$T$	Latent variables' score matrix, 潜在変数のスコア行列
USP	United States Pharmacopeia, 米国薬局方
VIP	Variable influence on projection, 射影における変数重要度
$V_m$	VIP score at the $m$ th wavenumber, $m$ 番目の波数における VIP 得点
$W$	Weighting matrix, 重み付け行列
$W(Z_g, \neg Z_g)$	The sum of affinities between the wavenumbers belonging to $Z_g$ and those not belonging to $Z_g$ , $Z_g$ に属する波数と $Z_g$ に属さない波数との類似度の 和
$W(Z_g)$	The sum of affinities between the wavenumbers belonging to $Z_g$ , $Z_g$ に属す る波数間の類似度の和
$W_{\max}$	The maximal value of $W_{\text{mov}}$ , $W_{\text{mov}}$ の最大値
$W_{\min}$	The minimal value of $W_{\text{mov}}$ , $W_{\text{mov}}$ の最小値
$W_{\text{mov}}$	Width of moving window, 移動枠の幅
$w_{\text{mov}}$	Width of moving window in a spectral interval スペクトル領域内の移動 枠の幅
$X$	Input variable matrix, 入力変数行列
$X_m$	The translated $X$ so that $\mathbf{x}_m$ becomes the origin, $\mathbf{x}_m$ が原点となるように変 換された $X$
$\mathbf{x}_m$	The $m$ th column vector of $X$ , $X$ の $m$ 番目の列ベクトル
$\bar{x}_m$	Average value of $\mathbf{x}_m$ , $\mathbf{x}_m$ の平均値
$\ \mathbf{x}_m - \mathbf{z}_g\ $	The Euclidean distance between $\mathbf{x}_m$ and $\mathbf{z}_g$ , $\mathbf{x}_m$ と $\mathbf{z}_g$ との間のユークリッド 距離
$\mathbf{x}_{mk}$	The $k$ th column vector of $X_m$ , $X_m$ の $k$ 番目の列ベクトル
$\mathbf{x}_{ml}$	The $l$ th column vector of $X_m$ , $X_m$ の $l$ 番目の列ベクトル
$x_{nm}$	Spectral intensity at the $m$ th wavenumber in the $n$ th sample, $n$ 番目のサン プルの $m$ 番目の波数におけるスペクトル強度

$\mathbf{y}$	Reference value vector of output variable, 出力変数の参照値ベクトル
$\hat{\mathbf{y}}$	Estimated value vector of output variable, 出力変数の推定値ベクトル
$\bar{\mathbf{y}}$	Average value of reference values of output variable in a validation set, バリデーションセットにおける出力変数の参照値の平均値
$y_n$	Reference value of output variable in the $n$ th sample, $n$ 番目のサンプルの出力変数の参照値
$\hat{y}_n$	Estimated value of output variable in the $n$ th sample, $n$ 番目のサンプルの出力変数の推定値
$Z_g$	The $g$ th group, $g$ 番目のグループ
$\mathbf{z}_g$	The mean vector of the $g$ th wavenumber group, $g$ 番目の波数グループの平均ベクトル
$\mathbf{z}_{gm}$	The mean vector of the group to which the $m$ th wavenumber is assigned, $m$ 番目の波数が割当てられたグループの平均ベクトル
$Z_i$	Positions of the wavenumbers included in the $i$ th spectral interval, $i$ 番目のスペクトル領域に含まれる波数の位置



## 第一章 スペクトル変動分割を用いた効率的な波数選択法の開発

### 1.1 序論

NIRS に基づく検量モデルを構築する際、推定精度及び頑健性を向上させるためには推定対象とする品質特性の変動を適切に反映する波数を選択することが重要である<sup>28)</sup>。そこで統計的な波数選択法が提案されている<sup>27), 34)</sup>。統計的な波数選択法における一つの方策は、個々の波数を品質特性との関連性に従って順位付けし、その順位の高いものから波数を一つ一つ選択するものである。この波数単位での変数選択法の代表例は PLS-beta 及び射影における変数重要度 (Variable influence on projection: VIP) である<sup>35)</sup>。もう一つの方策は、スペクトル全体を幾つかのスペクトル領域に分割し、各スペクトル領域を品質特性との関連性に従って順位付けし、その順位の高いものからスペクトル領域単位で波数を選択するものである。このスペクトル領域に基づく波数選択法の代表例は等幅のスペクトル領域を作製する iPLS である<sup>36)</sup>。元来、iPLS は推定精度の観点で重要なスペクトル領域を可視化する手段として開発されたが、波数選択法としても利用されている。隣接する波数におけるスペクトル強度は強く相関することが多く、それが特定のスペクトル領域における吸収ピークとして表れるため、スペクトル領域に着目した波数選択は有望な方策である。さらに iPLS をその他の手法と組合せた波数選択法が提案されており、その例として moving window PLS (MWPLS)<sup>37)</sup>、changeable size MWPLS (CSMWPLS)<sup>38)</sup>、searching combination MWPLS (SCMWPLS)<sup>38)</sup>、及び genetic algorithm-PLS (GA-PLS)<sup>39), 40)</sup>が挙げられる。これらの統計的な波数選択法は、最適な波数選択結果を得るために調節パラメータを最適化することが必要である。この調節パラメータを最適化するために膨大な数の検量モデルを評価するため、非常に大きな計算負荷を必要とする。

計算負荷は、波数選択法を日常的に運用する上で考慮すべき現実的な課題である。これは、検量モデルを開発する時点だけでなく製品ライフサイクルを通じて波数選択法が継続的に使用され得るためである。検量モデルを開発した時点において検量モデルが所望の推定精度を有していたとしても、日常生産における製造プロセスが変動し、それが原因となって推定精度が低下する恐れがある。そこで、製品のライフサイクルを通じて検量モデルをメンテナンスすること、つまり継続的にその推定精度を検証し、必要に応じて検量モデルを更新することが求められている<sup>31), 32), 41), 42)</sup>。検量モデルを更新する場合、その更新期間中は検量モデルを製造プロセス管理に使用することができない。この場合、検量モデルを用いる NIRS 以外の代替的な測定方法 (e.g., 高速液体クロマトグラフィー: HPLC) を一時的に採用することとなる。これは NIRS より非効率的な作業であり、NIRS のようにリアルタイムで品質を測定することが出来ない。従って、検量モデルを更新するために要する作業負荷及び期間は、医薬品を安定して製造及び供給する計画に影響を与えることが懸念される。検量モデルを更新する際、検量モデルの頑健性を向上させるために、改めて適切な波数を選択することが必要となることが多い。波数選択法の計算負荷が小さければ、より短い時間で波数選択を完

了することができる。その結果として、検量モデルを更新するために要する作業負荷を低減し、その期間を短縮することができる。

本章では、新規のスペクトル分割法に基づく効率的な波数選択法を提案する。提案するスペクトル分割法である SFD は、キャリブレーションセットにおける各波数点のスペクトル強度の標準偏差を算出し、スペクトル変動特性とする。そして、そのスペクトル変動特性の極小値でスペクトル全体を複数のスペクトル領域に分割する。SFD に基づく波数選択法は、各スペクトル領域に含まれる全ての波数を用いて構築した検量モデルの推定精度が高い順にスペクトル領域を選択する。提案する波数選択法を顆粒中の水分含量及び薬物（化合物 X）含量を推定する 2 種類の検量モデル構築に適用し、推定精度及び計算時間の観点で従来の波数選択法と比較評価した。

## 1.2 実験の部

### 1.2.1 材料

原薬（知的財産の観点で開示できない特定の化合物であるため、化合物 X とする）を含む顆粒（Daiichi-Sankyo, Japan）を測定サンプルとして用いた。キャリブレーションセット及びバリデーションセット用のサンプルとして使用した顆粒の製造スケールは、造粒工程において 4 kg/batch から 100 kg/batch 及び混合工程において 0.4 kg/batch から 500 kg/batch とした。なお、batch とは造粒機または混合機を一回運転して得られる顆粒をひとまとまりにした単位であり、各 batch の顆粒は均一な物質特性を有すると考えられる。造粒工程では、流動層造粒機を用いて原薬及び数種の添加剤を造粒した。使用した造粒機は 4 kg/batch スケールでは NFLO-5（Freund, Japan）、10 kg/batch スケールでは Aeromatic Fielder（GEA Pharma Systems, Belgium）、100 kg/batch スケールでは WSG-120（Powrex, Japan）または GPCG-120（Glatt, German）とした。混合工程では、混合機を用いてその顆粒と滑沢剤を混合した。使用した混合機は 0.4 kg/batch スケールでは S-3-S（Tsutsui Scientific Instruments, Japan）、2 kg/batch スケールでは TCV-10（Tokuju, Japan）、100 kg/batch から 300 kg/batch スケールでは PM-1000（Bohle, German）または TB-1200（Tanico, Japan）、500 kg/batch スケールでは PM-2000（Kotobuki, Japan）とした。造粒工程中及び混合工程後の顆粒をサンプリングした。

### 1.2.2 NIRS 測定

造粒工程では、造粒機缶体に設置した光ファイバプローブを介してフーリエ変換型 NIR 分光器 MPA（Bruker GmbH, Germany）または同等の Matrix-F（Bruker GmbH, Germany）を用いて NIRS 測定を実施し、造粒工程中 1 分毎に NIR スペクトルを取得した。混合工程では、サンプリングした顆粒約 0.2 g を専用バイアルに秤取し、MPA を用いてそのバイアルを測定することで NIR スペクトルを取得した。測定条件を Table 1 に示す。NIR スペクトルは OPUS 6.5 software（Bruker GmbH, Germany）を用いて記録した。記録する波数範囲を  $4000\text{ cm}^{-1}$  から  $12500\text{ cm}^{-1}$  と設定したが、NIR 分

光器を定期的にキャリブレーションする度に波数点を記録する位置が僅かに変動する。波数点の位置が僅かに異なる NIR スペクトルを解析するために、OPUS 6.5 software は各推定対象（水分含量及び薬物含量）の検量モデルで使用する NIR スペクトルが同じ波数点を有するように内挿法を用いて波数点の位置を調節している。従って、二つの推定対象の間で僅かに NIR スペクトルの波数点の数が異なる（水分含量の推定では 2201 波数点、薬物含量の推定では 2202 波数点）。

Table 1 Experimental conditions to prepare the calibration and validation sets used for constructing and validating PLS models.

	Water content estimation	Drug content estimation
NIRS measurement with the diffuse reflectance method		
Wavenumber range [cm <sup>-1</sup> ]	12500–4000	12500–4000
Resolution [cm <sup>-1</sup> ]	8	8
No. of wavenumber points	2201	2202
Integration time	8 times	64 times
No. of NIR spectra		
Calibration set	96 (13 batches)	64 (64 batches)
Validation set	58 (7 batches)	40 (40 batches)
Reference measurement		
Measurement method	LOD	HPLC
Range of value [%]		
Calibration set	1.1–17.0	67.7–130.7
Validation set	1.7–15.6	73.1–124.2

### 1.2.3 参照測定

造粒工程では、造粒中 10 分毎、スプレー終点、及び乾燥終点で流動層造粒機から顆粒をサンプリングした。サンプリングした顆粒の水分含量を HR73 (Mettler-Toledo, Japan) または同等の HR83 (Mettler-Toledo, Japan) を用いて乾燥減量 (LOD) 法により測定した。その水分含量値を対応する測定時点の NIR スペクトルと関連付けた。LOD 装置の技術情報から、LOD 法の測定精度は標準偏差 0.2% である。混合工程では、混合後の顆粒の薬物（化合物 X）含量を Waters Alliance 2695 Separations Module (Waters Corporation, US) を用いて HPLC 法により測定した。HPLC 法の添加回収法の結果から、HPLC 法の測定精度は標準偏差 0.4% である。

### 1.2.4 PLS

本項では PLS を概説する。PLS を適用するにあたり、予め入出力変数を平均中心化した。その入力変数行列（前処理後の NIR スペクトル） $\mathbf{X} \in \mathbf{R}^{N \times M}$  及び出力変数の参照値ベクトル（顆粒中の水分または薬物含量） $\mathbf{y} \in \mathbf{R}^N$  は以下のように分解される：

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E} \quad (1)$$

$$\mathbf{y} = \mathbf{Tb} + \mathbf{f} \quad (2)$$

ここで  $\mathbf{T} \in \mathbf{R}^{N \times K}$  はスコアベクトル  $\mathbf{t}_k \in \mathbf{R}^N$  ( $k=1, 2, \dots, K$ ) を列とする潜在変数の

スコア行列、 $\mathbf{P} \in \mathbf{R}^{M \times K}$  はローディングベクトル  $\mathbf{p}_k \in \mathbf{R}^M$  ( $k=1, 2, \dots, K$ ) を列とする  $\mathbf{X}$  のローディング行列、 $\mathbf{E} \in \mathbf{R}^{N \times M}$  は  $\mathbf{X}$  の残差行列、 $\mathbf{b} \in \mathbf{R}^M$  は  $\mathbf{y}$  の重み付けベクトル、 $\mathbf{f} \in \mathbf{R}^N$  は  $\mathbf{y}$  の残差ベクトル、 $N$  はサンプルの数、 $M$  は入力変数の数（波数点の数）、そして  $K$  は採用した潜在変数の数である。PLS を実施するアルゴリズムとして、非線形反復部分最小二乗法 (nonlinear iterative partial least squares: NIPALS) を採用した<sup>20)</sup>。1 番目から  $k-1$  番目のスコアベクトルを  $\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_{k-1}$ 、 $\mathbf{X}$  のローディングベクトルを  $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_{k-1}$ 、 $\mathbf{y}$  の重み付けベクトルの値を  $b_1, b_2, \dots, b_{k-1}$  とする。 $k$  番目の残差入力変数行列  $\mathbf{X}_k$  及び  $k$  番目の残差出力変数ベクトル  $\mathbf{y}_k$  は以下のように表現される:

$$\mathbf{X}_k = \mathbf{X}_{k-1} - \mathbf{t}_k \mathbf{p}_k^T \quad (3)$$

$$\mathbf{y}_k = \mathbf{y}_{k-1} - b_k \mathbf{t}_k^T \quad (4)$$

ここで、 $\mathbf{X}_0 = \mathbf{X}$  及び  $\mathbf{y}_0 = \mathbf{y}$  とする。 $k$  番目のスコアベクトル  $\mathbf{t}_k$  は以下の式で与えられる:

$$\mathbf{t}_k = \mathbf{X}_{k-1} \mathbf{w}_k \quad (5)$$

ここで  $\mathbf{w}_k \in \mathbf{R}^M$  は  $k$  番目の重み付けベクトルであり、 $\|\mathbf{w}_k\|=1$  の下で  $\mathbf{y}_{k-1}$  と  $\mathbf{t}_k$  との間の共分散を最大化するベクトルである:

$$\mathbf{w}_k = \frac{\mathbf{X}_{k-1}^T \mathbf{y}_{k-1}}{\|\mathbf{X}_{k-1}^T \mathbf{y}_{k-1}\|} \quad (6)$$

$\mathbf{X}$  の  $k$  番目のローディングベクトル  $\mathbf{p}_k$  及び  $\mathbf{b}$  の  $k$  番目の値  $b_k$  は以下の式で得られる:

$$\mathbf{p}_k = \frac{\mathbf{X}_{k-1}^T \mathbf{t}_k}{\mathbf{t}_k^T \mathbf{t}_k} \quad (7)$$

$$b_k = \frac{\mathbf{y}_{k-1}^T \mathbf{t}_k}{\mathbf{t}_k^T \mathbf{t}_k} \quad (8)$$

上記の手順を  $K$  回繰り返す。最適な  $K$  は一つ抜きクロスバリデーション (leave-one-out cross validation: LOOCV) を実施し、クロスバリデーションの標準誤差 (standard error of cross validation: SECV) を最小化するように選択される。SECV は LOOCV における推定誤差の指標であり、以下の式で算出される:

$$\text{SECV} = \sqrt{\frac{1}{N} \sum_{n=1}^N (\hat{y}_n - y_n)^2} \quad (9)$$

ここで、 $y_n$  は  $n$  番目のサンプルの出力変数の参照値及び  $\hat{y}_n$  は  $n$  番目のサンプルの出力変数の推定値である。 $K$  が大きくなるほど SECV は小さくなるが、検量モデルがキャリブレーションセットに対してオーバーフィッティングする危険性が高くなる。本検討では、キャリブレーションセットへのオーバーフィッティングを回避するために、経験的な知識に基づいて  $K$  を 1 から 20 の範囲で選択した。

### 1.2.5 キャリブレーションセット及びバリデーションセット並びに推定精度の指標

キャリブレーションセット及びバリデーションセットの概要を Table 1 に示す。造粒工程では、各ロットの造粒時間に応じて4から12時点でサンプリングを実施し、各サンプリング時点で1つのサンプルを作製した。水分含量を推定する検量モデルを構築するにあたり、そのキャリブレーションセット用のサンプルにおける水分含量を1.1%から17.0%の範囲で意図的に変動させた。検量モデルの推定精度を評価するにあたり、そのバリデーションセット用のサンプルにおける水分含量を1.7%から15.6%の範囲で意図的に変動させた。混合工程では、各ロットにおいて混合された顆粒をサンプリングし、1つのサンプルを作製した。薬物（化合物 X）含量を推定する検量モデルを構築するにあたり、そのキャリブレーションセット用のサンプルにおける薬物（化合物 X）含量を67.7%から130.7%の範囲で意図的に変動させた。その検量モデルの推定精度を評価するにあたり、そのバリデーションセット用のサンプルにおける薬物（化合物 X）含量を73.1%から124.2%の範囲で意図的に変動させた。キャリブレーションセット及びバリデーションセットを構成するサンプルは互いに独立させた。

キャリブレーションセットを用いて検量モデルを構築し、キャリブレーションの標準誤差（standard error of calibration: SEC）に基づいて検量精度を評価した。構築済の検量モデルを用いてバリデーションセットの品質特性（水分含量または薬物含量）を推定し、予測標準誤差（standard error of prediction: SEP）及び決定係数（ $R^2$ ）に基づいて推定精度を評価した。SEP と SEC との残差は検量モデルがキャリブレーションセットへオーバーフィッティングしている可能性の指標となる。SEC 及び SEP はそれぞれキャリブレーションセット及びバリデーションセットにおける全てのサンプルを用いて Eq. (9) の右辺と同じ式により計算される。 $R^2$  は以下の式で算出される:

$$R^2 = 1 - \frac{\sum_{n=1}^N (y_n - \hat{y}_n)^2}{\sum_{n=1}^N (y_n - \bar{y})^2} \quad (10)$$

ここで  $y_n$  及び  $\hat{y}_n$  はバリデーションセットの全てのサンプルについて評価する。また、 $\bar{y}$  はバリデーションセットにおける出力変数の参照値の平均値である。

### 1.2.6 スペクトル前処理

各推定対象について最適なスペクトル前処理を選択するために、以下の手法を比較評価した。

- 標準正規変量（standard normal variate: SNV）
- 一次微分

- 二次微分
- 一次微分+SNV
- 二次微分+SNV

SNV は各波数のスペクトル強度を標準化することで、多重散乱の影響を低減することができる<sup>25)</sup>。一次微分及び二次微分は NIR スペクトルにおける基線変動の影響を低減させるとともに、微小な吸収ピークを強調することができる。SNV と微分とを組み合わせることで、両手法の効果を相加的に得ることができる。本検討では、キャリブレーションセットにおける前処理後の NIR スペクトルの全波数を用いて検量モデルを構築し、SEP が最小となるスペクトル前処理を採用した。スペクトル前処理は OPUS 6.5 software を用いて適用した。なお、スペクトル前処理で適用する一次微分及び二次微分は、後述の SFD-PLS においてスペクトル変動特性の極小値を求めるための計算としては使用しない。以降、“スペクトル強度”は、全て前処理後の NIR スペクトルにおけるスペクトル強度を意味する。

### 1.2.7 計算時間

計算負荷の観点で波数選択法を評価するために、調節パラメータの最適化を含む各波数選択法の手順を一回実行するために要した計算時間を計測した。計算時間の計測には、同一のキャリブレーションセット及び計算機を用いた。計算時間はキャリブレーションセットのデータ量（サンプル数及び NIR スペクトルの波数点の数）、計算機の性能、及び波数選択アルゴリズムに依存する。特にキャリブレーションセットのデータ量が大きいほど、計算時間は指数関数的に長くなる。本検討では、キャリブレーションセットのサンプル数を 100 未満とした。これは計算時間を現実的に比較評価可能な範囲（数時間～数十時間）に収めるためである。現実的にはキャリブレーションセットのサンプル数は 1000 以上となる場合があり、その場合は日単位ないし年単位の計算時間を要する。

### 1.3 波数選択法

NIRS に基づいて製剤品質を管理するにあたり、検量モデルの推定精度及び頑健性を向上させることが重要である。NIR スペクトルの全波数を用いて検量モデルを構築すると最適な推定精度及び頑健性が得られない可能性があるため、検量モデルを構築する上で波数選択は事実上必須の作業である。

本検討では、以下の波数選択法を用いて検量モデルを構築し、その推定精度及び計算負荷を比較評価した。

従来法:

- 1) PLS-beta （波数単位で評価）
- 2) VIP （波数単位で評価）
- 3) iPLS （スペクトル領域単位で評価）

提案法 (スペクトル領域単位で評価) :

- 4) SFD-PLS
- 5) 移動平均 SFD-PLS (moving averaged SFD-PLS: MASFD-PLS)
- 6) SFD 面積-単回帰分析法 (SFD area-simple linear regression: SFDA-SLR)

波数選択法の参照として、波数選択を実施せず NIR スペクトルの全波数を用いる検量モデル (PLS-All) を別途評価した。波数選択法の適用には、ソフトウェアとして MATLAB<sup>®</sup> 2014a software (The MathWorks, US) 及び計算機として HP ProBook 4320s (Hewlett-Packard, US. OS: Windows 7 Professional 32 bit, CPU: Intel<sup>®</sup> Core i5 2.67 GHz, RAM: 4.00 GB) を用いた。なお、本計算機の性能は一般的な家庭用計算機で得られる程度の水準である。提案する波数選択法は一般的な家庭用計算機を用いて実用可能であることを、計算時間の観点で検証する。

### 1.3.1 PLS-beta

従来法である PLS-beta は PLS に基づく検量モデルの回帰係数の絶対値が高い順に波数を選択する<sup>35)</sup>。出力変数の推定値ベクトル ( $\hat{y} \in \mathbf{R}^N$ ) 及び回帰係数ベクトル ( $\beta_{\text{PLS}} \in \mathbf{R}^{M_{\text{all}}}$ ) は以下の式で得られる:

$$\hat{y} = T\mathbf{b} = X\mathbf{W}(\mathbf{P}^T\mathbf{W})^{-1}\mathbf{b} = X\beta_{\text{PLS}} \quad (11)$$

$$\beta_{\text{PLS}} = \mathbf{W}(\mathbf{P}^T\mathbf{W})^{-1}\mathbf{b} \quad (12)$$

ここで  $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_K] \in \mathbf{R}^{M_{\text{all}} \times K}$  は重み付け行列、 $M_{\text{all}}$  はスペクトル全体における入力変数の数 (波数点の数) である。PLS-beta の調節パラメータは  $M$  である。

PLS-beta に基づく波数選択手順を以下に示す:

- 1) スペクトルの全波数を用いて検量モデルを構築し、 $\beta_{\text{PLS}}$  を算出する。
- 2)  $\beta_{\text{PLS}}$  の要素の絶対値が大きい順に波数を並べ替え、 $M=1$  とする。
- 3) 最初の  $M$  個の波数を用いて検量モデルを構築し、SECV を算出する。
- 4)  $M=M+1$  とし、 $M=M_{\text{all}}-1$  となるときまで step 3 に進む ( $M=M_{\text{all}}$  である場合は PLS-All に相当するので省略する)。
- 5) SECV が最小となる  $M$  を選択し、 $\beta_{\text{PLS}}$  の要素の絶対値が大きい順に  $M$  個の波数を採用する。

### 1.3.2 VIP

従来法である VIP は個々の入力変数が出力変数の推定値に与える影響の大きさの指標である VIP 得点が大きい順に波数を選択する<sup>20)</sup>。 $m$  番目の入力変数の VIP 得点 ( $V_m$ ) は次のように定義される:

$$V_m = \sqrt{\frac{M_{\text{all}} \sum_{k=1}^K (w_{km}^2 b_k^2 t_k^T t_k)}{\sum_{k=1}^K (b_k^2 t_k^T t_k)}} \quad (13)$$

ここで  $w_{km}$  は  $k$  番目の重み付けベクトル ( $w_k$ ) の  $m$  番目の要素である。VIP の調節パラメータは  $M$  である。

VIP に基づく波数選択手順を以下に示す:

- 1) スペクトルの全波数を用いて検量モデルを構築し、 $V_m$  を算出する。
- 2)  $V_m$  が大きい順に波数を並べ替え、 $M=1$  とする。
- 3) 最初の  $M$  個の波数を用いて検量モデルを構築し、SECV を算出する。
- 4)  $M=M+1$  とし、 $M=M_{\text{all}}-1$  となるときまで step 3 に進む ( $M=M_{\text{all}}$  である場合は PLS-All に相当するので省略する)。
- 5) SECV が最小となる  $M$  を選択し、 $V_m$  が大きい順に  $M$  個の波数を採用する。

### 1.3.3 iPLS

従来法である iPLS はスペクトル全体を等幅のスペクトル領域に分割し、各スペクトル領域に含まれる全ての波数を用いて検量モデルを構築する。そして、その推定精度が高い順にスペクトル領域を選択する<sup>36)</sup>。iPLS の概念図を Fig. 3 に示す。

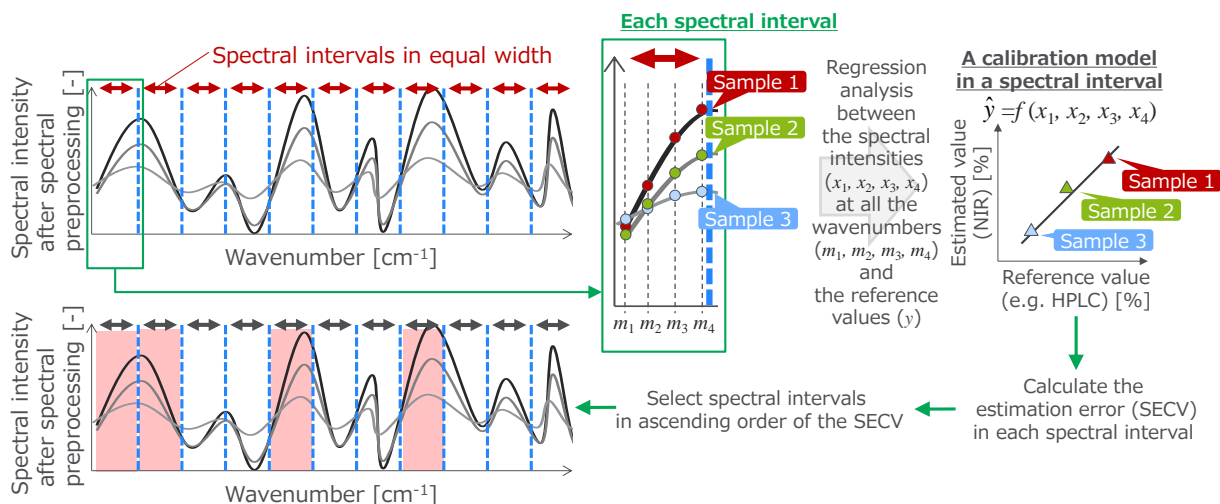


Fig. 3 An illustrative example of iPLS. In Fig.3, it is assumed that a spectral interval is composed of four wavenumbers. [-] : Spectral intensity after spectral processing is dimensionless.

本検討では、分割するスペクトル領域の数 ( $I_{\text{div}}$ ) を  $I_{\text{min}}=3$  から  $I_{\text{max}}=100$  の範囲で設定した。ここで  $I_{\text{min}}$  及び  $I_{\text{max}}$  はそれぞれ  $I_{\text{div}}$  の最小値及び最大値である。iPLS の調節パラメータは  $I_{\text{div}}$  及び選択するスペクトル領域の数 ( $I_{\text{sel}}$ ) である。



iPLS に基づく波数選択手順を以下に示す:

- 1)  $I_{div}=I_{min}$  とする。
- 2) スペクトル全体を  $I_{div}$  個の等幅のスペクトル領域に分割する。
- 3) 各スペクトル領域に含まれる全ての波数を用いて検量モデルを構築し、SECV を算出する。
- 4) SECV が小さい順にスペクトル領域を並べ換え、 $I_{sel}=1$  とする。
- 5) 最初の  $I_{sel}$  個のスペクトル領域に含まれる全ての波数を用いて検量モデルを構築し、SECV を算出する。
- 6)  $I_{sel}=I_{sel}+1$  とし、 $I_{sel}=I_{div}-1$  となるときまで step 5 に進む ( $I_{sel}=I_{div}$  である場合は PLS-All に相当するので省略する)。
- 7)  $I_{div}=I_{div}+1$  とし、 $I_{div}=I_{max}$  となるときまで step 2 に進む。
- 8) SECV を最小とする  $[I_{sel}, I_{div}]$  の組合せを選択し、最初の  $I_{sel}$  個のスペクトル領域に含まれる波数を採用する。

### 1.3.4 SFD-PLS (提案法)

本項以降では本章で提案する新規手法について述べる。

NIRS において化学物質の濃度が変動すると、その吸収帯に対応する特定の吸収ピーク (スペクトル領域におけるスペクトル強度) が変動する。しかし、複数の物質特性に由来する吸収ピークが重なり合うため、目的とする物質特性に対応するスペクトル領域を分光学的な知識のみに基づいて正確に特定することは困難である。もしスペクトル全体を各吸収ピークに対応するスペクトル領域に分割することができれば、等幅のスペクトル領域を用いる iPLS よりも正確に物質特性の変動を反映する波数を選択できると考えられた。ここで、仮に一つの NIR スペクトルにおけるスペクトル強度の極小値をスペクトル領域の境界として用いた場合、Fig. 4 に示すように同一の吸収ピークに対応する極小値の位置は各 NIR スペクトルに応じて異なる。従って、NIR スペクトル毎に異なるスペクトル領域の境界を検出することになる。また、吸収ピークの形状はそれが由来する物質特性に応じて異なり、その数も膨大であることから、各ピーク形状に対し分布関数を仮定して吸収ピークを認識することも困難である。

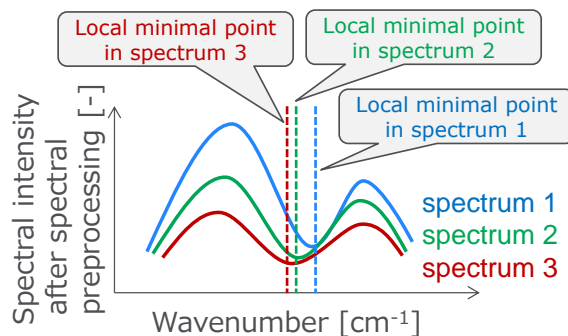


Fig. 4 An illustrative example of local minimal point of the spectral intensities in each NIR spectrum.

[-] : Spectral intensity after spectral processing is dimensionless.

SFD は各吸収ピークに対応したスペクトル領域を検出することを目的とする。吸収ピークの境界では前処理後 NIR スペクトルのスペクトル強度の変動が極小となると考えられる。Fig. 5 に示すように、SFD はキャリブレーションセットに含まれる全ての NIR スペクトルを対象として各波数点のスペクトル強度の標準偏差を算出し、スペクトル変動特性とする。そして、そのスペクトル変動特性の極小値でスペクトル全体を複数のスペクトル領域に分割する。 $m$  番目の波数におけるスペクトル強度 ( $\mathbf{x}_m \in \mathbf{R}^N$ ) の標準偏差 ( $s_m$ ) は以下の式で算出される:

$$s_m = \sqrt{\frac{1}{N-1} \sum_{n=1}^N (x_{nm} - \bar{x}_m)^2} \quad (14)$$

ここで  $x_{nm}$  は  $n$  番目のサンプルの  $m$  番目の波数におけるスペクトル強度、 $\bar{x}_m$  は  $\mathbf{x}_m$  の平均値である。SFD はスペクトル変動特性の極小点をスペクトル領域の境界として用いる。この極小点で分割されたスペクトル領域は各吸収ピークに対応することが期待される。なお、SFD はスペクトル前処理後の NIR スペクトルに対して適用する (1.2.6 スペクトル前処理 参照)。

提案する SFD-PLS は各スペクトル領域に含まれる全ての波数を用いて検量モデルを構築し、その推定精度が高い順にスペクトル領域を選択する。SFD-PLS の調節パラメータは  $I_{sel}$  のみである。iPLS と異なり、SFD によって  $I_{div}$  が一意的に定まる。

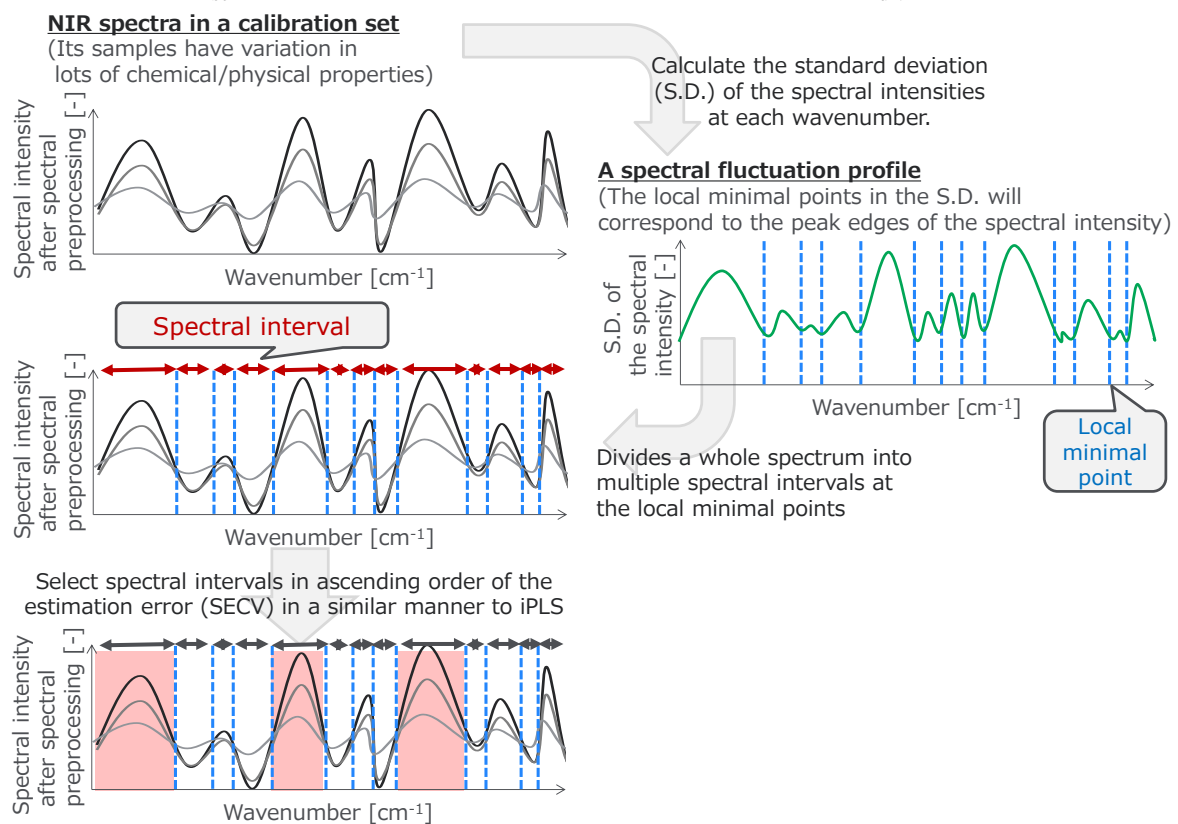


Fig. 5 An illustrative example of SFD-PLS. [-] : Spectral intensity after spectral processing is dimensionless.

SFD-PLS に基づく波数選択手順を以下に示す:

- 1) キャリブレーションセットにおける各波数のスペクトル強度の標準偏差を算出し、スペクトル変動特性とする。
- 2) スペクトル全体をスペクトル変動特性の極小点で分割し、 $I_{div}$  個のスペクトル領域とする。
- 3) 各スペクトル領域に含まれる全ての波数を用いて検量モデルを構築し、SECV を算出する。
- 4) SECV が小さい順にスペクトル領域を並べ換え、 $I_{sel}=1$  とする。
- 5) 最初の  $I_{sel}$  個のスペクトル領域に含まれる全ての波数を用いて検量モデルを構築し、SECV を算出する。
- 6)  $I_{sel}=I_{sel}+1$  とし、 $I_{sel}=I_{div}-1$  となるときまで step 5 に進む ( $I_{sel}=I_{div}$  である場合は PLS-All に相当するので省略する)。
- 7) SECV を最小とする  $I_{sel}$  を選択し、最初の  $I_{sel}$  個のスペクトル領域に含まれる波数を採用する。

### 1.3.5 MASFD-PLS (提案法)

SFD はスペクトル変動の大きさを考慮せず、スペクトル変動特性 (キャリブレーションセットとして使用する NIR スペクトルにおける各波数点のスペクトル強度の標準偏差) の全ての極小点を用いてスペクトル全域を分割する。この時、Fig. 6 に示すように測定ノイズに由来する微小なスペクトル変動が生じた場合であってもスペクトル変動特性に極小点が生成されるため、化学的な情報とは無関係のスペクトル領域が生成される可能性がある。そこで測定ノイズがスペクトル変動特性に与える影響を低減する方策として、スペクトル変動特性にスムージング処理を適用する。一般的なスムージング処理の 1 つとして移動平均法がある。本検討では SFD-PLS と移動平均法を組合せ、その波数選択法を MASFD-PLS と名付けた。MASFD-PLS の調節パラメータは  $I_{sel}$  及び移動平均法で用いる移動枠の幅 ( $W_{mov}$ ) である。本検討では、 $W_{mov}$  を  $W_{min}=3$  から  $W_{max}=9$  の範囲で移動枠の刻み幅  $\Delta W_{mov}=2$  として設定した。ここで、 $W_{min}$  及び  $W_{max}$  はそれぞれ  $W_{mov}$  の最小値及び最大値である。MASFD-PLS は SFD-PLS よりも少ない  $I_{div}$  が得られるため、評価する波数の組合せが少なくなり、計算負荷を低減できると期待される。

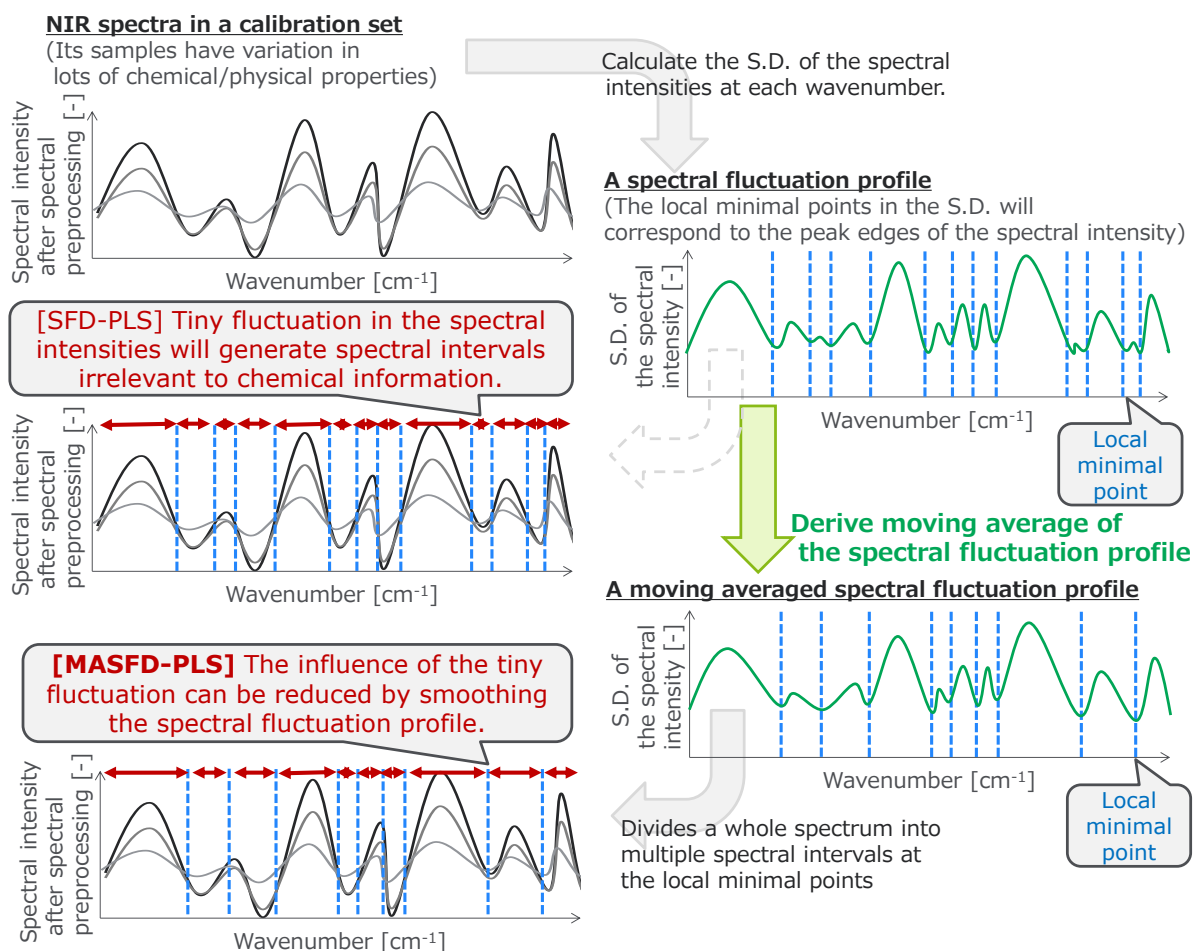


Fig. 6 An illustrative example of the moving averaged spectral fluctuation profile in MASFD-PLS.

[-] : Spectral intensity after spectral processing is dimensionless.

MASFD-PLS に基づく波数選択手順を以下に示す:

- 1) キャリブレーションセットにおける各波数のスペクトル強度の標準偏差を算出し、スペクトル変動特性とする。
- 2)  $W_{mov}=W_{min}$  とする。
- 3)  $W_{mov}$  を移動枠の幅としてスペクトル変動特性に移動平均法を適用することで、移動平均スペクトル変動特性を作製する。
- 4) スペクトル全体を移動平均スペクトル変動特性の極小点で分割し、 $I_{div}$  個のスペクトル領域とする。
- 5) 各スペクトル領域に含まれる全ての波数を用いて検量モデルを構築し、SECV を算出する。
- 6) SECV が小さい順にスペクトル領域を並べ替え、 $I_{sel}=1$  とする。
- 7) 最初の  $I_{sel}$  個のスペクトル領域に含まれる全ての波数を用いて検量モデルを構築し、SECV を算出する。

- 8)  $I_{sel}=I_{sel}+1$  とし、 $I_{sel}=I_{div}-1$  となるときまで step 7 に進む ( $I_{sel}=I_{div}$  である場合は PLS-All に相当するので省略する)。
- 9)  $W_{mov}=W_{mov}+\Delta W_{mov}$  とし、 $W_{mov}=W_{max}$  となるときまで step 3 に進む。
- 10) SECV を最小とする  $[W_{mov}, I_{sel}]$  の組合せを選択し、最初の  $I_{sel}$  個のスペクトル領域に含まれる波数を採用する。

### 1.3.6 SFDA-SLR (提案法)

SFD は各吸収ピークに対応するスペクトル領域を生成することを意図した手法である。ここで各スペクトル領域におけるスペクトル強度の曲線下面積はスペクトル情報の指標となり得る。この面積を SFD 面積と名付けた。なお、SFD 面積は前処理後の NIR スペクトルにおけるスペクトル強度を用いて算出する。各波数点におけるスペクトル強度の代わりに、各スペクトル領域における SFD 面積を入力変数として用いることができる。SFD 面積を入力変数として用いた場合、スペクトル全体に含まれる入力変数の数が少なくなるため、波数選択の計算負荷を低減できる。SFD 面積に基づく波数選択法を SFDA-SLR と名付けた。SFDA-SLR の調節パラメータは  $I_{sel}$  のみである。SFDA-SLR では SFD-PLS とは異なり、各スペクトル領域におけるモデル構築に SLR を用いた。これは各スペクトル領域に含まれる入力変数 (SFD 面積) の数が 1 つであるためである。

SFDA-SLR に基づく波数選択手順を以下に示す:

- 1) キャリブレーションセットにおける各波数のスペクトル強度の標準偏差を算出し、スペクトル変動特性とする。
- 2) スペクトル全体をスペクトル変動特性の極小点で分割し、 $I_{div}$  個のスペクトル領域とする。
- 3) 各スペクトル領域に含まれる全ての波数点におけるスペクトル強度の和を算出し、SFD 面積とする。
- 4) 各スペクトル領域につき、SFD 面積を入力変数として用いて SLR に基づく検量モデルを構築し、SECV を算出する。
- 5) SECV が小さい順にスペクトル領域を並べ換え、 $I_{sel}=1$  とする。
- 6) 最初の  $I_{sel}$  個のスペクトル領域に含まれる SFD 面積を入力変数として用いて PLS に基づく検量モデルを構築し、SECV を算出する。
- 7)  $I_{sel}=I_{sel}+1$  とし、 $I_{sel}=I_{div}-1$  となるときまで step 6 に進む ( $I_{sel}=I_{div}$  である場合は PLS-All に相当するので省略する)。
- 8) SECV を最小とする  $I_{sel}$  を選択し、最初の  $I_{sel}$  個のスペクトル領域を採用する (SFD 面積を入力変数として用いるため、波数ではなくスペクトル領域の選択となる)。

## 1.4 結果及び考察

### 1.4.1 スペクトル前処理

顆粒中の水分含量及び薬物（化合物 X）含量を推定する 2 種類の検量モデルを構築するにあたり、スペクトル前処理を最適化した。SNV、一次微分、二次微分、一次微分+SNV、及び二次微分+SNV の 5 つのスペクトル前処理を推定精度の観点で比較した。キャリブレーションセットにおける前処理後の NIR スペクトルの全波数を用いて検量モデルを構築し、その推定精度を比較するためにバリデーションセットを用いて SEP を算出した。結果を Table 2 及び Table 3 に示す。水分含量推定の場合是一次微分、薬物（化合物 X）含量推定の場合是一次微分+SNV を用いた場合に SEP が最小となった。従って、これらのスペクトル前処理を採用し、以降の波数選択法を比較評価する検討に用いることとした。

Table 2 Comparison of spectral preprocessing methods in the water content estimation.

	$K$	SECV [%]	SEC [%]	SEP [%]	$R^2$
SNV	19	0.41	0.16	0.56	0.95
First derivative	18	0.30	0.16	0.36	0.98
Second derivative	20	0.62	0.29	0.79	0.92
First derivative + SNV	5	0.94	0.79	1.18	0.84
Second derivative + SNV	15	1.07	0.66	1.15	0.83

Table 3 Comparison of spectral preprocessing methods in the drug content estimation.

	$K$	SECV [%]	SEC [%]	SEP [%]	$R^2$
SNV	20	2.34	0.13	3.65	0.85
First derivative	20	2.16	0.13	3.80	0.85
Second derivative	20	3.51	0.08	3.43	0.85
First derivative + SNV	20	2.42	0.13	2.93	0.88
Second derivative + SNV	20	3.79	0.07	3.68	0.82

### 1.4.2 スペクトル変動分割

本章では、新規スペクトル分割法として SFD 及び MASFD を提案する。これらの提案法では、スペクトル変動特性に基づいてスペクトル全体を複数のスペクトル領域に分割する。顆粒中の水分含量及び薬物（化合物 X）含量を推定する 2 種類の検量モデルを構築するにあたり、SFD 及び MASFD をキャリブレーションセットの NIR スペクトルに適用した。それにより得られたスペクトル変動特性及び移動平均スペクトル変動特性を Fig. 7 及び Fig. 8 に示す。また、これらのスペクトル変動特性及び移動平均スペクトル変動特性における極小点の数を Table 4 及び Table 5 に示す。これらの結果から、MASFD はスペクトル変動特性の移動平均を算出することで、 $I_{div}$  を低減できることが検証された。MASFD-PLS は評価する波数組合せの数が SFD-PLS より少なくなるため、計算負荷を低減できると期待される。

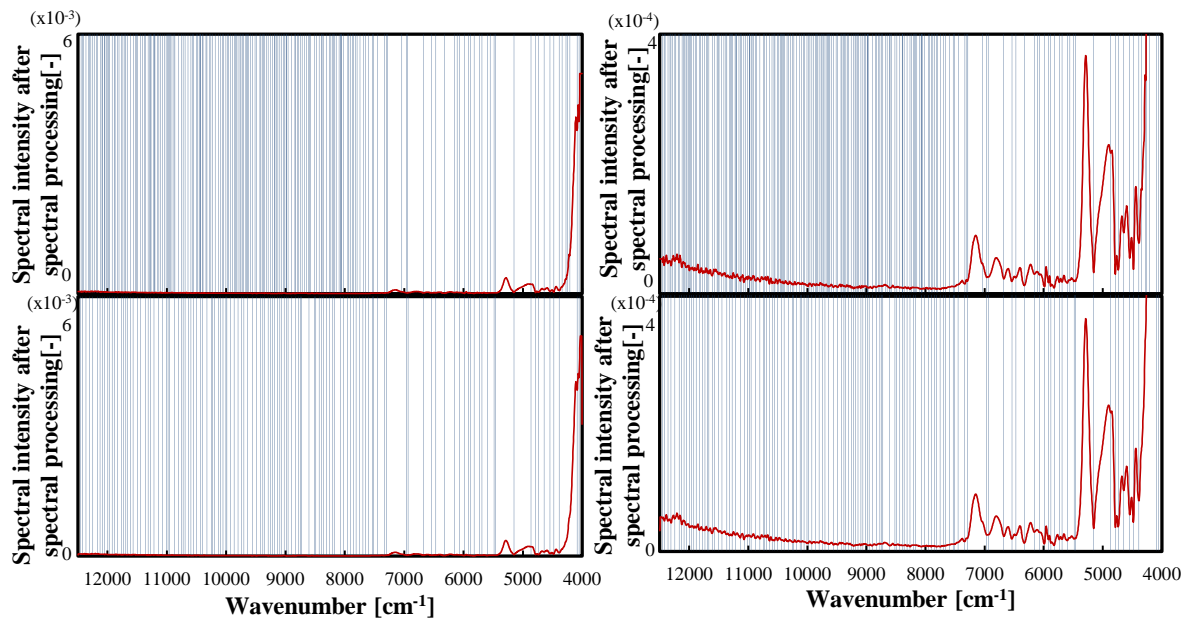


Fig. 7 The spectral fluctuation profiles in the water content estimation. Vertical lines indicate the local minimal points. (Left–Top) the spectral fluctuation profile in SFD, (Left–Bottom) the moving averaged spectral fluctuation profile in MASFD, (Right–Top) enlarged view of the spectral fluctuation profile in SFD, and (Right–Bottom) enlarged view of the moving averaged spectral fluctuation profile in MASFD. [-] : Spectral intensity after spectral processing is dimensionless.

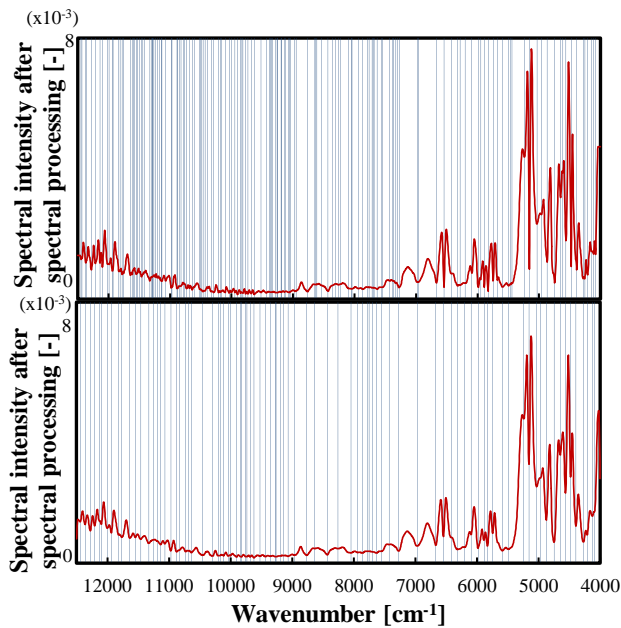


Fig. 8 The spectral fluctuation profiles in the drug content estimation. Vertical lines indicate the local minimal points. (Top) the spectral fluctuation profile in SFD and (Bottom) the moving averaged spectral fluctuation profile in MASFD. [-] : Spectral intensity after spectral processing is dimensionless.

Table 4 The numbers of local minimal points (#LocalMin), divided spectral intervals ( $I_{div}$ ), and the number of evaluated combinations of input variables (#eval) in SFD-PLS and MASFD-PLS with different moving window sizes ( $W_{mov}$ ) in the water content estimation.

	#LocalMin	$I_{div}$	#eval
SFD ( $W_{mov}=1$ )	206	207	206
MASFD ( $W_{mov}=3$ )	167	168	167
MASFD ( $W_{mov}=5$ ) <sup>a</sup>	129	130	129
MASFD ( $W_{mov}=7$ )	115	116	115
MASFD ( $W_{mov}=9$ )	108	109	108

a: adopted in the result of MASFD-PLS

Table 5 The numbers of local minimal points (#LocalMin), divided spectral intervals ( $I_{div}$ ), and the number of evaluated combinations of input variables (#eval) in SFD-PLS and MASFD-PLS with different moving window sizes ( $W_{mov}$ ) in the drug content estimation.

	#LocalMin	$I_{div}$	#eval
SFD ( $W_{mov}=1$ )	155	156	155
MASFD ( $W_{mov}=3$ )	125	126	125
MASFD ( $W_{mov}=5$ )	106	107	106
MASFD ( $W_{mov}=7$ )	96	97	96
MASFD ( $W_{mov}=9$ ) <sup>a</sup>	88	89	88

a: adopted in the result of MASFD-PLS

### 1.4.3 波数選択及び検量モデル構築

本項では、「1.3 波数選択法」項に示した 6 つの波数選択法を顆粒中の水分含量及び薬物（化合物 X）含量を推定する検量モデルの構築に適用し、それらの波数選択法を推定精度の観点で比較する。

#### 1.4.3.1 水分含量推定

選択された調節パラメータ及び推定精度を Table 6 に示し、選択された波数を Fig. 9 に図示する。水分は  $11800\text{ cm}^{-1}$ 、 $10200\text{ cm}^{-1}$ 、 $8600\text{ cm}^{-1}$ 、 $7000\text{ cm}^{-1}$ 、 $5600\text{ cm}^{-1}$ 、及び  $5200\text{ cm}^{-1}$  付近に NIR 吸収帯を有する<sup>43)</sup>。Fig. 9 (b, c) に示すように、これらの吸収帯のうち  $7000\text{ cm}^{-1}$  及び  $5200\text{ cm}^{-1}$  付近の吸収帯は、キャリブレーションセットにおける  $7150\text{ cm}^{-1}$  及び  $5300\text{ cm}^{-1}$  付近の主要な吸収ピークと一致している。従って、これらの主要な吸収ピークは自由水に由来すると考えられた。キャリブレーションセットにおける  $5000\text{ cm}^{-1}$  付近の双峰性ピークも水分と関連付けることができるが、この双峰性ピークは自由水の影響だけではなく原薬（水和物である化合物 X）における結合水の影響をも受けている。一般に、水分に関連する波数領域におけるスペクトル強度は水分含量だけでなく薬物含量や顆粒の物理的特性等の水分とは無関係な特性の影響を受ける。水分と無関係な特性の変動が推定精度に与える影響を最小化するために、水分に関連する波数領域だけでなく水分と無関係な波数領域を選択することが水分含量を推定する検量モデル構築に有用である<sup>12)</sup>。

PLS-beta は、2201 波数点のうち 308 波数点を選択した。VIP は 2201 波数点のうち 108 波数点を選択した。iPLS は NIR スペクトル全体を 38 個のスペクトル領域に分



割し、そのうち 5 個のスペクトル領域を選択した。SFD-PLS は NIR スペクトル全体を 207 個のスペクトル領域に分割し、そのうち 57 個のスペクトル領域を選択した。MASFD-PLS は NIR スペクトル全体を 130 個のスペクトル領域に分割し、そのうち 44 個のスペクトル領域を選択した。SFDA-SLR は NIR スペクトル全体を 207 個のスペクトル領域に分割し、そのうち 96 個のスペクトル領域を選択した。

全ての波数選択法において、選択された波数は  $7150\text{ cm}^{-1}$  または  $5300\text{ cm}^{-1}$  付近の水分に基づく主要なピーク領域を含んでいたため、水分含量の変動を反映する波数領域を採用したと考えられる。また、全ての波数選択法において、選択された波数は水分と無関係の波数領域をも含んでいた。

PLS-beta 及び VIP において、選択された波数は顕著な測定ノイズを示す  $4000\text{ cm}^{-1}$  付近の波数領域を含んでいた。従って、検量モデルの推定精度は測定ノイズの影響を受けやすいと考えられる。一方、iPLS、SFD-PLS、MASFD-PLS、及び SFDA-SLR において、選択された波数は  $4000\text{ cm}^{-1}$  付近の波数領域を除外していたため、検量モデルの推定精度は測定ノイズに対してより頑健となることが期待される。

全ての波数選択法の中で、提案する SFD-PLS 及び MASFD-PLS は SEP 及び SEP と SEC との残差を最小化した。この結果は、SFD-PLS 及び MASFD-PLS は推定精度を最も向上させるとともに、キャリブレーションセットにオーバーフィッティングする危険性を最も低減したことを示している。これは SFD-PLS 及び MASFD-PLS が水分と関連する波数領域及び水分と無関係であるが推定精度の向上に有用な波数領域の両方を最も適切に選択した結果と推察される。一方で、SFDA-SLR は SFD-PLS よりも大きい SEP（低い推定精度）を示した。この原因は以下のように考えられる。Fig. 4 に示すように、吸収ピークの境界付近におけるスペクトル強度は、隣接する吸収ピークの影響を受ける。従って、同一の吸収ピークにおけるスペクトル強度と推定対象とする品質特性の変動との相関関係は、そのスペクトル領域内の波数位置によって異なる。これを適切に反映するために、波数単位で回帰係数に重みを付けることが必要である。SFD-PLS は各波数のスペクトル強度を入力変数として用いており、波数単位で回帰係数に重みを付ける。一方、SFDA-SLR は各スペクトル領域における複数のスペクトル強度の和である SFD 面積を入力変数として用いており、各 SFD 面積に回帰係数の重みを付ける。つまり、SFDA-SLR は各スペクトル領域における全ての波数のスペクトル強度に等しい重みを付けていることに等しい。この違いが原因となり、SFDA-SLR は SFD-PLS よりも大きい SEP を示したと考えられる。

Table 6 Comparison of wavenumber selection methods in the water content estimation (spectral preprocessing method: first derivative).

	$M$ ( $I_{sel}/I_{div}$ )	$K$	#eval <sup>a</sup>	Computation time [h]	SECV [%]	SEC [%]	SEP [%]	$R^2$
PLS-All	2201	18	1	- <sup>b</sup>	0.30	0.16	0.36	0.98
PLS-beta	308	20	2200	26.78	0.24	0.16	0.39	0.98
VIP	108	20	2200	26.92	0.27	0.18	0.33	0.98
iPLS	285 (5/38)	4	4949	56.13	0.24	0.22	0.32	0.99
SFD-PLS	978 (57/207)	6	206	1.24	0.25	0.19	0.29	0.99
MASFD-PLS	998 (44/130) $W=5^c$	6	519 ( $W_{mov}=3,5,7,9)^d$	7.40	0.25	0.19	0.29	0.99
SFDA-SLR	96 <sup>e</sup> (96/207)	9	206	0.07	0.24	0.18	0.36	0.98

a: the number of evaluated combinations of input variables. b: wavenumber selection was not performed. c: adopted as the final model. d: four  $W_{mov}$  were evaluated. e: SFD areas were used as input variables.

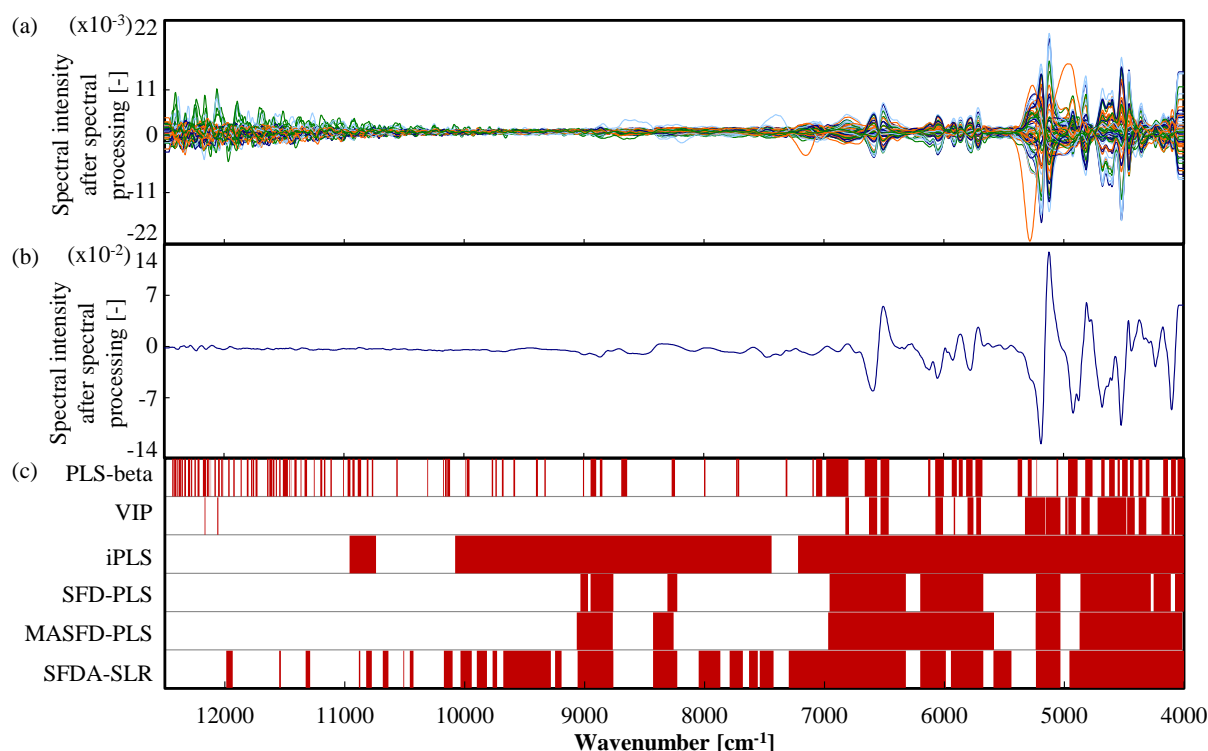


Fig. 9 Wavenumber selection results in the water content estimation (spectral preprocessing method: first derivative). (a) the preprocessed NIR spectra of the calibration set. (b) enlarged view of the preprocessed NIR spectra in the calibration set. (c) the selected wavenumbers (shaded regions). [-] : Spectral intensity after spectral processing is dimensionless.

#### 1.4.3.2 薬物含量推定

選択された調節パラメータ及び推定精度を Table 7 に示し、選択された波数を Fig. 10 に図示する。Fig. 10 (b) に示すように、薬物 (化合物 X) は  $6500\text{ cm}^{-1}$ 、 $5150\text{ cm}^{-1}$

<sup>1</sup>、4800 cm<sup>-1</sup>、4500 cm<sup>-1</sup>、及び 4100 cm<sup>-1</sup> 付近において主要な吸収ピークを示す。Fig. 10 (a, b) に示すように、これらの薬物（化合物 X）の吸収ピークのうち、6500 cm<sup>-1</sup> 及び 5150 cm<sup>-1</sup> 付近の吸収ピークは、キャリブレーションセットにおける主要な吸収ピークと一致している。7150 cm<sup>-1</sup>、5300 cm<sup>-1</sup>、及び 5000 cm<sup>-1</sup> 付近の波数領域において顕著な変動を示す NIR スペクトルは、高い水分含量を有するキャリブレーション用のサンプルに相当する。従って、「1.4.3.1 水分含量推定」項に記載の通り、7150 cm<sup>-1</sup> 及び 5300 cm<sup>-1</sup> 付近のスペクトル強度は自由水を反映し、5000 cm<sup>-1</sup> 付近のスペクトル強度は自由水及び薬物（化合物 X）における結合水を反映していると考えられる。水分含量の変動を予め検量モデルに組み込み、水分含量の変動が推定精度に与える影響を低減させる目的で、この高水分含量のサンプルを意図的にキャリブレーションセット用のサンプルとして採用した。水分含量の場合と同様に、薬物（化合物 X）に関連する波数領域におけるスペクトル強度は薬物（化合物 X）含量だけでなく水分含量や顆粒の物理的特性等の薬物（化合物 X）とは無関係な特性の影響を受ける。薬物（化合物 X）と無関係な特性の変動が推定精度に与える影響を最小化するために、薬物（化合物 X）に関連する波数領域だけでなく薬物（化合物 X）と無関係な波数領域を選択することが薬物（化合物 X）含量を推定する検量モデル構築に有用であると考えられた。

PLS-beta は、2202 波数点のうち 601 波数点を選択した。VIP は 2202 波数点のうち 343 波数点を選択した。iPLS は NIR スペクトル全体を 38 個のスペクトル領域に分割し、そのうち 27 個のスペクトル領域を選択した。SFD-PLS は NIR スペクトル全体を 156 個のスペクトル領域に分割し、そのうち 28 個のスペクトル領域を選択した。MASFD-PLS は NIR スペクトル全体を 89 個のスペクトル領域に分割し、そのうち 26 個のスペクトル領域を選択した。SFDA-SLR は NIR スペクトル全体を 156 個のスペクトル領域に分割し、そのうち 75 個のスペクトル領域を選択した。

全ての波数選択法において、選択された波数は 6500 cm<sup>-1</sup> または 5150 cm<sup>-1</sup> 付近の薬物（化合物 X）に基づく主要なピーク領域を含んでいたため、薬物（化合物 X）含量の変動を反映する波数領域を採用したと考えられる。

PLS-beta、VIP、及び SFDA-SLR において、選択された波数は測定ノイズが相対的に大きい 12000 cm<sup>-1</sup> 付近の波数領域を含んでいた。従って、得られる検量モデルの推定精度は測定ノイズの影響を受けやすいと考えられる。一方、iPLS、SFD-PLS、及び MASFD-PLS において、選択された波数は 12000 cm<sup>-1</sup> 付近の波数領域を除外していたため、得られる検量モデルの推定精度は測定ノイズに対してより頑健となることが期待される。

水分含量の推定における結果と同様に、全ての波数選択法の中で提案する SFD-PLS は SEP 及び SEP と SEC との残差を最小化した。また、MASFD-PLS は SFD-PLS と同程度の SEP 及び SEP と SEC との残差を得た。この結果は、SFD-PLS 及び MASFD-PLS は推定精度を最も向上させるとともに、キャリブレーションセットにオーバーフ

フィッティングする危険性を最も低減したことを示している。これは SFD-PLS 及び MASFD-PLS が薬物（化合物 X）と関連する波数領域及び薬物（化合物 X）と無関係であるが推定精度の向上に有用な波数領域の両方を最も適切に選択した結果と推察される。また、SFDA-SLR は SFD-PLS より大きい SEP（低い推定精度）を示したが、この結果も水分含量推定における結果と一致している。

水分含量及び薬物（化合物 X）含量の推定において得られたこれらの結果から、提案する SFD-PLS 及び MASFD-PLS はその他の手法よりも推定精度及びキャリブレーションセットにオーバーフィッティングする危険性の両方の観点で優れていることが示された。

Table 7 Comparison of wavenumber selection methods in the drug content estimation (spectral preprocessing method: first derivative + SNV).

	$M$ ( $I_{sel}/I_{div}$ )	$K$	#eval <sup>a</sup>	Computation time [h]	SECV [%]	SEC [%]	SEP [%]	$R^2$
PLS-All	2202	20	1	- <sup>b</sup>	2.42	0.13	2.93	0.88
PLS-beta	601	20	2201	17.24	0.94	0.08	3.13	0.87
VIP	343	14	2201	17.08	1.56	1.00	2.46	0.92
iPLS	1575 (27/38)	20	4949	37.61	1.45	0.25	2.53	0.92
SFD-PLS	649 (28/156)	13	155	1.46	1.57	1.03	1.81	0.96
MASFD-PLS	756 (26/89) $W=9^c$	13	415 ( $W_{mov}=3,5,7,9^d$ )	3.62	1.52	1.01	1.81	0.96
SFDA-SLR	75 <sup>e</sup> (75/156)	20	155	0.03	1.68	0.60	2.51	0.92

a: the number of evaluated combinations of input variables. b: wavenumber selection was not performed. c: adopted as the final model. d: four  $W_{mov}$  were evaluated. e: SFD areas were used as input variables.

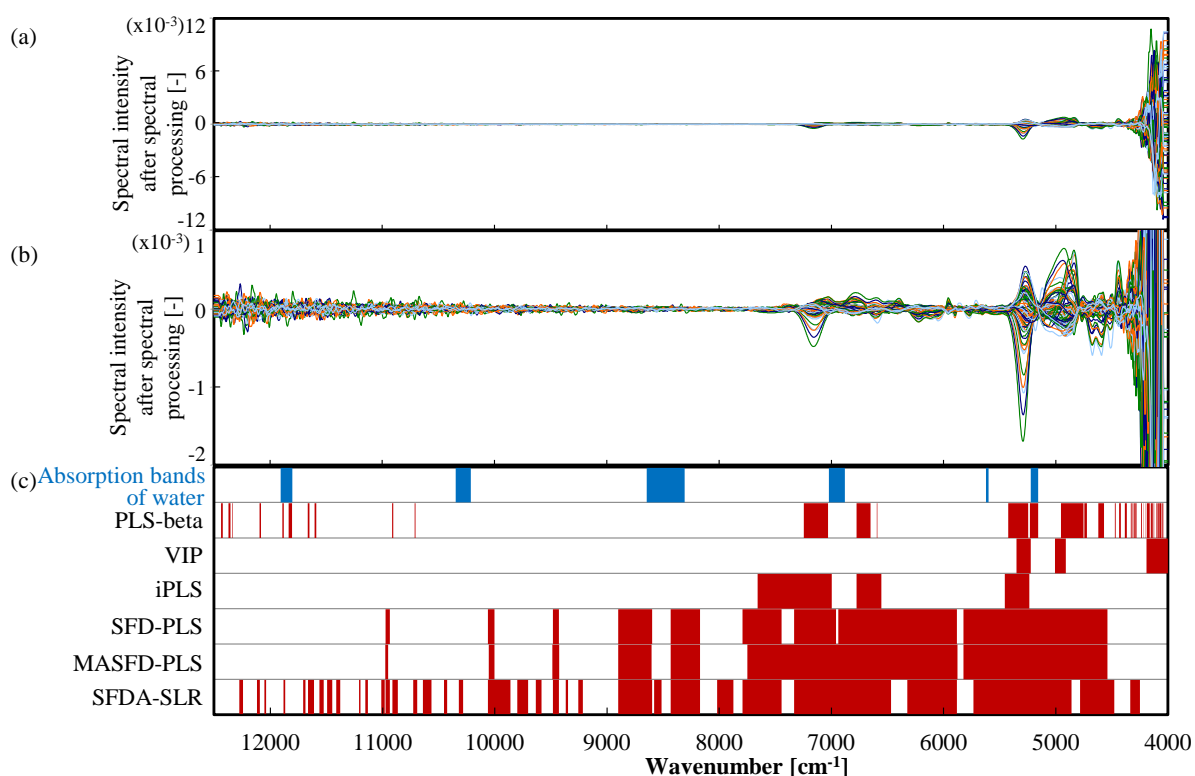


Fig. 10 Wavenumber selection results in the drug content estimation (spectral preprocessing method: first derivative + SNV). (a) the preprocessed NIR spectra of the calibration set. (b) the preprocessed NIR spectrum of the drug substance. (c) the selected wavenumbers (shaded regions). [-] : Spectral intensity after spectral processing is dimensionless.

#### 1.4.4 計算時間

本項では、「1.3 波数選択法」項に示した6つの波数選択法を顆粒中の水分含量及び薬物（化合物 X）含量を推定する2種類の検量モデルの開発に適用した際の計算負荷を比較する。計算負荷は、各波数選択法が要した計算時間によって評価した。計算時間を Table 6 及び Table 7 に示す。計算時間は、水分含量及び薬物（化合物 X）含量を推定する事例で同様の傾向がみられた。

各波数選択法における計算時間は、SFDA-SLR を除くと、評価した波数の組合せの数にほぼ比例した。提案する SFD-PLS 及び MASFD-PLS は従来法である PLS-beta、VIP、及び iPLS よりも短い計算時間で波数選択を完了させた。特に SFD-PLS で要した計算時間は、PLS-beta 及び VIP と比較して 1/10 以下、iPLS と比較して 1/25 以下であった。各移動枠の幅  $W_{mov}$  (=3、5、7、または9) における MASFD-PLS の計算時間は SFD-PLS よりも短かったが、MASFD-PLS は4つの  $W_{mov}$  (=3、5、7、及び9) を評価したため、最終的に SFD-PLS よりも長い計算時間を要した。以上の結果から、推定精度及び計算時間の両方の観点で提案する SFD-PLS 及び MASFD-PLS は従来法よりも優れていることが示された。

SFDA-SLR は最も短い計算時間で波数選択を完了させた。これは SFDA-SLR が

入力変数として使用する SFD 面積の数が、その他の波数選択法が入力変数として使用する波数点の数の 1/10 以下となったためである。しかし、「1.4.3 波数選択及び検量モデル構築」項で示した通り、SFDA-SLR は SFD-PLS よりも低い推定精度を示した。

以上をまとめると、推定精度及び計算負荷の観点で、SFD-PLS 及び MASFD-PLS は本検討で評価した波数選択法の中で最良の手法である。

## 1.5 小括

本章では、NIRS に基づく検量モデルの開発に向けて新規のスペクトル分割法に基づく波数選択法を提案した。従来の波数選択法である iPLS はスペクトル全体を等幅のスペクトル領域に分割し、そのスペクトル領域単位で波数を選択する。しかし、等幅のスペクトル領域では推定精度を向上させる上で有用な波数を柔軟に選択することができない。一方、提案するスペクトル分割法 SFD はスペクトル全体を各吸収ピークに対応するスペクトル領域に分割することを意図したものである。吸収ピークの境界ではスペクトル強度の変動が極小となると考えられる。そこで SFD は、まずキャリブレーションセットにおける各波数のスペクトル強度の標準偏差を算出し、それをスペクトル変動特性とする。そして、そのスペクトル変動特性の極小点においてスペクトル全体を複数のスペクトル領域に分割する。これによって、推定精度を向上させる上で有用な波数をより適切に選択できることが期待される。SFD に基づく波数選択法 SFD-PLS は、各スペクトル領域におけるスペクトル強度を用いて検量モデルを構築し、その推定精度が高い順にスペクトル領域を選択する。

事例研究として、提案する波数選択法を用いて顆粒中の水分含量及び薬物（化合物 X）含量を推定する 2 種類の検量モデルを構築し、推定精度及び計算時間の観点で従来の波数選択法と比較評価した。両事例において、提案する SFD-PLS は従来法である PLS-beta、VIP、及び iPLS よりも高い推定精度をより小さな計算負荷で達成した。なお、SFD-PLS は一般的な家庭用計算機を用いて 1~2 時間で波数選択を完了した。これらの結果は、SFD-PLS が推定対象とする品質特性に関連する波数領域を最も正確にかつ効率的に選択したことを示唆している。さらに SFD-PLS に移動平均法を組合せた MASFD-PLS は推定精度を損なわずに分割するスペクトル領域の数を低減させた。従って、SFD-PLS 及び MASFD-PLS は従来法よりも推定精度及び計算負荷の観点で優れている。

今回、異なる 2 つの推定対象に対して同様の結果が得られたことから、広範な推定対象に対してこれらの提案法を適用できると考えられる。本検討では顆粒のみを測定対象としたが、一般に NIR 吸収帯に応じて吸収ピークが生じ、その吸収ピークが重なり合う等の NIR スペクトルの特徴は、測定対象に依らず同様である。従って、顆粒以外の様々な測定対象に対してもこれらの提案法を適用できるだろう。これらの提案法に基づく波数選択を実行するプログラムを作成しておけば、そのプログラムを実行することで自動的に波数選択を実施できる。従って、波数選択を実施する担当者

に依らず同一の結果を得られる。また、推定対象や測定対象に依らず、同一のプログラムを用いて波数選択を実施できる。以上から、これらの提案法は推定精度及び計算負荷の観点で優れているだけでなく、波数選択結果が個人の能力に依存しない波数選択法であり、多様な推定対象及び測定対象に適用できる。

## 第二章 相関に基づくクラスタリングを用いた効率的な波数選択法の開発

### 2.1 序論

第一章ではスペクトル領域に基づく波数選択法を提案したが、第二章ではさらにスペクトル領域間のスペクトル強度の相関を考慮した波数選択法を提案する。

第一章で論じた通り、推定精度及び頑健性の高い検量モデルを開発する上で、波数選択は重要な役割を担う。本章では、波数をまとめることで波数グループを作製し、その波数グループ単位で波数を選択する手法を検討する。iPLS等のスペクトル領域に基づく手法では、各スペクトル領域に含まれる全ての波数を波数グループと見なすことができる。スペクトル領域とは別に、波数間の類似性に基づいて個々の波数を波数グループにクラスタリングすることも可能である。クラスタリングに基づく手法として、NCSCに基づく変数選択が提案されている<sup>44)</sup>。なお、NCSCは当初検量モデル開発におけるキャリブレーション用のサンプルをクラスタリングするために開発されたものである<sup>45),46)</sup>。NCSCは、波数間のスペクトル強度の相関に基づいて波数を波数グループにクラスタリングする。NIRSにおいて化学物質の濃度が変動すると単一の波数だけではなくその吸収帯に対応する複数の波数におけるスペクトル強度が変動する。また、一つの化学物質は複数の吸収帯を有する。これらのNIRスペクトルの特徴を考慮すると、スペクトル強度の相関に基づくクラスタリングは波数を適切にグループ化する有望な手段と考えられる。これらのスペクトル領域に基づく手法及びクラスタリングに基づく手法は推定精度を改善させることに成功している。しかし、これらの手法は最適な波数選択結果を得るために調節パラメータを最適化することが必要である。この調節パラメータを最適化するために膨大な数の検量モデルを評価するため、非常に大きな計算負荷を必要とする。一般に、より多くの波数の組合せを評価すればより高い推定精度が得られる可能性が高まるため、計算負荷と推定精度は相反する傾向を示すが、波数選択法のアルゴリズム次第で計算負荷及び推定精度の両方を改善することが可能である。

品質特性を推定する上で有用なスペクトル領域(e.g., 個々の吸収ピークに対応するスペクトル領域)は様々な幅を有する。しかし従来法であるiPLSは等幅のスペクトル領域を作製するため、スペクトル領域を柔軟に選択することができない。これとは別の従来法であるSCMWPLSはMWPLS及びCSMWPLSを組合せた手法であり、作製するスペクトル領域の幅を調節可能である。しかし、MWPLSは移動枠を用いて各波数点を近似的に評価し、その評価結果に基づいてスペクトル領域を作製する。さらにCSMWPLSはMWPLSにより作製した各スペクトル領域から唯一つの副スペクトル領域(一つのスペクトル領域に内包されるスペクトル領域)を抽出する。しかしMWPLSにによって作製された各スペクトル領域には推定精度を向上させるために有用なスペクトル領域が複数存在する可能性がある。こうしたSCMWPLSのアルゴリズムでは、適切なスペクトル領域を選択し、推定精度を向上させることが出来ない可能性がある。第一章では、小さな計算負荷で推定精度を改善する波数選択法として、ス



ペクトル領域に基づく手法 **SFD-PLS** を提案した。**SFD-PLS** は各吸収ピークに対応するスペクトル領域を作製することを意図したものであり、従来法である **iPLS** よりも小さな計算負荷で高い推定精度を達成した。しかし、そのスペクトル領域の数は膨大であり、より効率的に高い推定精度を得る観点で改善の余地がある。

第二章では、各変動因子に由来する複数の吸収ピークに対応するスペクトル領域グループを作製することを意図した手法を提案する。本提案法では、**SFD** を用いてスペクトル領域を作製し、**NCSC** 用いてそのスペクトル領域をクラスタリングすることで数個のスペクトル領域グループを作製する。このスペクトル領域のクラスタリングに基づく波数選択法を **SFD-NCSC-PLS** と名付ける。**SFD-NCSC-PLS** は各スペクトル領域グループに含まれる全ての波数を用いて検量モデルを構築し、その推定精度が高い順にスペクトル領域グループを選択する。推定精度及び計算負荷の観点から **SFD-NCSC-PLS** の有用性を示すため、従来法である **iPLS** 及び **SCMWPLS** 並びに **SFD** または **NCSC** のどちらか一方を用いた波数選択法と比較した。事例研究として、これらの波数選択法を顆粒中の水分含量及び薬物（化合物 **X**）含量を推定する 2 種類の検量モデル構築に適用し、推定精度及び計算時間の観点で評価した。

## 2.2 実験の部

第二章では、第一章と同じサンプル、測定法、生データ、回帰手法、及び評価指標を用いた。

### 2.2.1 材料

「1.2.1 材料」項を参照のこと。

### 2.2.2 NIRS 測定

「1.2.2 NIRS 測定」項を参照のこと。

### 2.2.3 参照測定

「1.2.3 参照測定」項を参照のこと。

### 2.2.4 PLS

「1.2.4 PLS」項を参照のこと。

### 2.2.5 キャリブレーションセット及びバリデーションセット並びに推定精度の指標

「1.2.5 キャリブレーションセット及びバリデーションセット並びに推定精度の指標」項を参照のこと。

## 2.2.6 スペクトル前処理

第一章において「1.4.1 スペクトル前処理」項で論じた通り、水分含量及び薬物（化合物 X）含量を推定する検量モデル構築に向けたスペクトル前処理を最適化済である。SNV、一次微分、二次微分、一次微分+SNV、及び二次微分+SNV を推定精度の観点で比較した結果、水分含量の推定において一次微分、薬物（化合物 X）含量の推定において一次微分+SNV を採用した場合に SEP が最小となったため、これらのスペクトル前処理を採用した。

## 2.2.7 計算時間

「1.2.7 計算時間」項を参照のこと。

## 2.3 波数選択法

一般に、化学物質は倍音及び結合音に応じた複数の NIR 吸収帯を有しているため、複数の吸収ピーク（スペクトル領域におけるスペクトル強度）は類似した挙動を示す。ただし、異なる化学物質に由来する吸収ピークが重なり合うため、スペクトル強度が類似した挙動を示すスペクトル領域の位置が不明瞭となることが多い。もし各吸収ピークに対応するスペクトル領域を作製でき、さらに吸収ピークが変動する挙動の類似性に従ってそのスペクトル領域をスペクトル領域グループにクラスタリングできれば、それらのスペクトル領域グループは品質特性の変動を反映する波数を効率的に検出する上で有用であろう。提案する SFD-NCSC-PLS は各変動因子に由来する複数の吸収ピークに対応するスペクトル領域グループを作製し、そのスペクトル領域グループ単位で波数を選択することで、検量モデルの推定精度を効率的に向上させることを意図したものである。

本章では、SFD-NCSC-PLS をその他の波数グループに基づく手法と比較する。スペクトル領域に基づく手法については、各スペクトル領域における波数を波数グループと定義する。波数クラスタリングに基づく手法については、スペクトル強度の類似性に基づいてクラスタリングした波数を波数グループと定義する。領域クラスタリングに基づく手法は、領域面積の類似性に基づいてクラスタリングした各スペクトル領域グループに含まれる全ての波数を波数グループと定義する。領域面積は以下のように算出する：

$$A_{ni} = \sum_{m \in Z_i} |x_{nm}| \quad (15)$$

ここで  $A_{ni}$  は  $n$  番目のサンプルにおける  $i$  番目のスペクトル領域の領域面積、 $x_{nm}$  は  $n$  番目のサンプルにおける  $m$  番目の波数における入力変数（スペクトル強度）、そして  $Z_i$  は  $i$  番目のスペクトル領域に含まれる波数の位置である。

顆粒中の水分含量及び薬物（化合物 X）含量を推定する検量モデルを構築するにあたり、以下の波数選択法を評価した。括弧内に各手法の簡単な説明を併記する。

#### スペクトル領域に基づく手法:

- 1) iPLS (等幅のスペクトル領域)
- 2) SCMWPLS (MWPLS により作製したスペクトル領域内における可変幅の副スペクトル領域)
- 3) SFD-PLS (SFD に基づくスペクトル領域)

#### 波数クラスタリングに基づく手法:

- 4)  $k$ -means-PLS (波数を  $k$ -means 法によりクラスタリング)
- 5) NCSC-PLS (波数を NCSC によりクラスタリング)

#### 領域クラスタリングに基づく手法:

- 6) interval- $k$ -means-PLS (等幅のスペクトル領域を  $k$ -means 法によりクラスタリング)
- 7) interval-NCSC-PLS (等幅のスペクトル領域を NCSC によりクラスタリング)
- 8) SFD- $k$ -means-PLS (SFD により作製したスペクトル領域を  $k$ -means 法によりクラスタリング)
- 9) SFD-NCSC-PLS (SFD により作製したスペクトル領域を NCSC によりクラスタリング;提案法)

波数選択法の参照として、波数選択を実施せず NIR スペクトルの全波数を用いた検量モデル (PLS-AII) を別途評価した。波数選択の適用には、ソフトウェアとして MATLAB<sup>®</sup> R2014a software (The MathWorks, US) 及び計算機として HP ProBook 4320s (Hewlett-Packard, US. OS: Windows 7 Professional 32 bit, CPU: Intel<sup>®</sup> Core i5 2.67 GHz, RAM: 4.00 GB) を用いた。なお、本計算機の性能は一般的な家庭用計算機で得られる程度の水準である。提案する波数選択法は一般的な家庭用計算機を用いて実用可能であることを、計算時間の観点で検証する。

本章では、分割するスペクトル領域の数を  $I_{\text{div}}$ 、作製する波数グループの数を  $G_{\text{gen}}$ 、そして選択する波数グループの数を  $G_{\text{sel}}$  とする。幾つかの手法では、 $I_{\text{div}}$  及び  $G_{\text{gen}}$  が等しくなる。

### 2.3.1 iPLS

従来法である iPLS はスペクトル全体を等幅のスペクトル領域に分割し、各スペクトル領域に含まれる全ての波数を波数グループと定義する。そして、各波数グループに含まれる全ての波数を用いて検量モデルを構築し、その SECV が小さい順に波数グループを選択する<sup>36)</sup>。 $I_{\text{div}}$  は  $I_{\text{min}}=3$  から  $I_{\text{max}}=100$  の範囲で最適化した。ここで  $I_{\text{min}}$  及び  $I_{\text{max}}$  はそれぞれ  $I_{\text{div}}$  の最小値及び最大値である。 $G_{\text{gen}}$  は  $I_{\text{div}}$  と等しいので、iPLS における調節パラメータは  $I_{\text{div}}$  及び  $G_{\text{sel}}$  である。 $[I_{\text{div}}, G_{\text{sel}}]$  の組合せは、SECV が最小となるように選択した。

### 2.3.2 SCMWPLS

従来法である SCMWPLS は MWPLS 及び CSMWPLS を組合せて波数を選択する<sup>38)</sup>。MWPLS は、予め設定した幅 ( $W_{\text{mov}}=2W_{\text{half}}+1$ ) の移動枠を作製し、その移動枠の位置をスペクトル全体に渡って移動させる。ここで移動枠の位置を各移動枠の中央の波数と定義する。各位置において、移動枠に含まれる全ての波数を用いて検量モデルを構築する。予め設定した数 ( $N_p$ ) の位置を SECV の小さいものから選択する。そして、選択された位置における  $N_p$  点の波数につき、連続する波数をまとめてスペクトル領域と定義する。本検討では、移動枠の幅は  $W_{\text{half}}=10$  に固定した。これは、移動枠の幅が吸収帯または化学物質の吸収ピークを検出できるだけの狭い枠幅であれば、移動枠の幅は MWPLS の結果に有意な影響を与えないためである<sup>37)</sup>。また、Arakawa らが提案する手順<sup>39)</sup>に従い  $N_p=M_{\text{all}}/2$  とした。仮に  $N_p$  を 1 から  $M_{\text{all}}$  まで等の網羅的な範囲で最適化したならば、SCMWPLS は甚大な計算時間を必要とするため、非現実的である。次に、MWPLS により作製した各スペクトル領域に対して CSMWPLS を適用することで、有用な副スペクトル領域 (一つのスペクトル領域に内包されるスペクトル領域) を選択する。CSMWPLS は、MWPLS により作製した各スペクトル領域内で可変幅の移動枠を移動させ、各移動枠に含まれる全ての波数を用いて検量モデルを構築する。そして、SECV が最小となる移動枠を副スペクトル領域として採用する。副スペクトル領域の組合せは SECV が最小となるように選択した。

### 2.3.3 SFD-PLS

本章では、SFD-PLS を提案法に対する比較対象として用いる。

NIRS において化学物質の濃度が変動すると、その吸収帯に相当する特定の吸収ピーク (スペクトル領域におけるスペクトル強度) が変動する。吸収ピークの境界ではスペクトル強度の変動が極小となると考えられる。SFD はこの吸収ピークに対応するスペクトル領域を検出することを意図したスペクトル分割法である。SFD はキャリブレーションセットにおける各波数のスペクトル強度の標準偏差を算出してスペクトル変動特性とし、そのスペクトル変動特性の極小点でスペクトル全体を複数のスペクトル領域に分割する。SFD-PLS は各スペクトル領域に含まれる全ての波数を波数グループと定義し、各波数グループに含まれる全ての波数を用いて検量モデルを構築し、その SECV が小さい順に波数グループを選択する。 $I_{\text{div}}$  は SFD によって自動的に定まる。また、 $I_{\text{div}}$  は  $G_{\text{gen}}$  に等しいので、SFD-PLS の調節パラメータは  $G_{\text{sel}}$  のみである。 $G_{\text{sel}}$  は SECV が最小となるように選択した。

### 2.3.4 k-means-PLS

本章では、k-means-PLS を提案法に対する比較対象として用いる。

k-means 法はスペクトル強度間の距離に基づいて波数を波数グループにクラスタリングする。手順は以下の 4 段階である: 1) 全ての波数を無作為にグループ分けし、

波数グループを作製する、2) 各波数と波数グループの中心との距離を算出する、3) その距離が最小となるように各波数を波数グループに割り当てる、4) 全ての波数の割り当てが収束するか、割り当て回数 ( $Q$ ) が予め設定した最大値 ( $Q_{\max}$ ) に到達するまで、steps 2 及び 3 を繰り返す。Step 1 で実施する最初の波数グループの分け方が最終的に得られる波数グループに影響を与える。そこで、初期化回数 ( $P$ ) が予め設定した最大値 ( $P_{\max}$ ) に到達するまで steps 1 から 4 を繰り返し、最も適切にクラスタリングされた結果(各波数と波数グループの中心との距離の和が最小になる結果)を採用する。本検討では、 $P_{\max}=10$  及び  $Q_{\max}=100$  とした。

$k$ -means-PLS は各波数グループに含まれる全ての波数を用いて検量モデルを構築し、その SECV が小さい順に波数グループを選択する。 $k$ -means-PLS における調節パラメータは  $G_{\text{gen}}$  及び  $G_{\text{sel}}$  である。 $G_{\text{gen}}$  は  $G_{\min}=3$  から  $G_{\max}=7$  の範囲で最適化した。ここで  $G_{\min}$  及び  $G_{\max}$  は  $G_{\text{gen}}$  の最小値及び最大値である。 $[G_{\text{gen}}, G_{\text{sel}}]$  の組合せは SECV が最小となるように選択した。

### 2.3.5 NCSC-PLS

本章では、NCSC-PLS を提案法に対する比較対象として用いる。

NCSC は NC 法及び SC 法により波数を波数グループにクラスタリングする<sup>44), 47)</sup>。NC 法は各波数におけるスペクトル強度 (入力変数) の相関に基づいて波数間の類似度行列  $S \in \mathbf{R}^{M \times M}$  を算出する。NC 法は以下の手順で実施する:

- 1)  $S=0$  及び  $m=1$  とする。ここで  $0$  は零行列である。
- 2)  $X_m = X - x_m \mathbf{1}^T$  とする。ここで  $X_m \in \mathbf{R}^{N \times M}$  は  $m$  番目の波数の入力変数ベクトル ( $x_m \in \mathbf{R}^N$ ) が原点となるように変換された  $X$ 、 $X \in \mathbf{R}^{N \times M}$  はキャリブレーションセットにおける入力変数行列、そして  $\mathbf{1} = [1, 1, \dots, 1]^T \in \mathbf{R}^M$  である。
- 3) 全ての  $k$  及び  $l$  ( $k \neq l$ ,  $k \in [1, 2, \dots, M]$ 、及び  $l \in [1, 2, \dots, M]$ ) の組み合わせについて  $|C_{k,l}|$  を算出する。ここで  $C_{k,l}$  は  $x_{mk} \in \mathbf{R}^N$  及び  $x_{ml} \in \mathbf{R}^N$  との間の相関係数、 $x_{mk}$  及び  $x_{ml}$  は  $X_m$  の  $k$  番目及び  $l$  番目の列ベクトルである。
- 4) もし  $|C_{k,l}| \geq R_{\text{thres}}$  であれば  $S(k, l) = S(k, l) + 1$  とする。ここで  $R_{\text{thres}}$  は予め設定する相関係数の閾値、 $S(k, l)$  は  $S$  の  $(k, l)$  番目の要素である。
- 5)  $m=m+1$  とし、 $m=M_{\text{all}}$  となるときまで step 2 に進む。

SC 法は波数間の類似度行列である  $S$  に基づいて波数を波数グループにクラスタリングする。SC 法の 1 つである min-max 分割法は、グループ間の類似度を最小化すると同時にグループ内の類似度を最大化する。これらの類似度は以下のように定義される:

$$W(Z_g, \neg Z_g) = \sum_{a \in Z_g, b \notin Z_g} S(a, b) \quad (16)$$

$$W(Z_g) = \sum_{a \in Z_g, a' \in Z_g} S(a, a') \quad (17)$$

ここで  $W(Z_g, \neg Z_g)$  は  $g$  番目のグループ ( $Z_g$ ) に属する波数と  $Z_g$  に属さない波数との間の類似度の総和、 $W(Z_g)$  は  $g$  番目のグループ ( $Z_g$ ) に属する波数同士の類似度の総和である。

Min-max 分割法の目的関数は以下の式で与えられる:

$$\operatorname{argmin} \sum_{g=1}^{G_{\text{gen}}} \frac{W(Z_g, \neg Z_g)}{W(Z_g)} \quad (18)$$

この目的関数は Ng et al. が提案する手法で解くことができる<sup>48)</sup>。

NCSC-PLS は各波数グループに含まれる全ての波数を用いて検量モデルを構築し、その推定精度が高い順に波数グループを選択する。NCSC-PLS における調節パラメータは  $R_{\text{thres}}$ 、 $G_{\text{gen}}$ 、及び  $G_{\text{sel}}$  である。 $G_{\text{gen}}$  は  $G_{\text{min}}=3$  から  $G_{\text{max}}=7$  の範囲で最適化し、 $R_{\text{thres}}$  は  $R_{\text{thres1}}=0.90$  または  $R_{\text{thres2}}=0.99$  とした。 $[R_{\text{thres}}, G_{\text{gen}}, G_{\text{sel}}]$  の組合せは SECV が最小となるように選択した。

### 2.3.6 Interval-k-means-PLS

本章では、interval-k-means-PLS を提案法に対する比較対象として用いる。

Interval-k-means 法は、以下のようにスペクトル領域を作製する。まず、スペクトル全体を等幅のスペクトル領域に分割する。次に、Eq. (15) を用いて領域面積を算出する。そして、k-means 法を領域面積に適用することで、スペクトル領域をユークリッド距離に基づくスペクトル領域グループにクラスタリングする。各スペクトル領域グループに含まれる全ての波数を波数グループと定義する。

Interval-k-means-PLS は各波数グループに含まれる全ての波数を用いて検量モデルを構築し、その SECV が小さい順に波数グループを選択する。Interval-k-means-PLS における調節パラメータは  $I_{\text{div}}$ 、 $G_{\text{gen}}$ 、及び  $G_{\text{sel}}$  である。 $I_{\text{div}}$  及び  $G_{\text{gen}}$  は、それぞれ  $I_{\text{min}}=3$  から  $I_{\text{max}}=100$  の範囲及び  $G_{\text{min}}=3$  から  $G_{\text{max}}=7$  の範囲で最適化した。 $[I_{\text{div}}, G_{\text{gen}}, G_{\text{sel}}]$  の組合せは SECV が最小となるように選択した。

### 2.3.7 Interval-NCSC-PLS

本章では、interval-NCSC-PLS を提案法に対する比較対象として用いる。

Interval-NCSC 法は、interval-k-means 法における k-means 法の代わりに NCSC を用いてスペクトル領域グループ、つまり波数グループを作製する。Interval-NCSC-PLS は各波数グループに含まれる全ての波数を用いて検量モデルを構築し、その SECV が小さい順に波数グループを選択する。Interval-NCSC-PLS における調節パラメータは  $I_{\text{div}}$ 、 $R_{\text{thres}}$ 、 $G_{\text{gen}}$ 、及び  $G_{\text{sel}}$  である。これらの調節パラメータは NCSC-PLS 及び interval-k-means-PLS と同様の範囲で最適化した。 $[I_{\text{div}}, R_{\text{thres}}, G_{\text{gen}}, G_{\text{sel}}]$  の組合せは SECV が最小となるように選択した。

### 2.3.8 SFD-*k*-means-PLS

本章では、SFD-*k*-means-PLS を提案法に対する比較対象として用いる。

SFD-*k*-means 法は、SFD 及び *k*-means 法を用いてスペクトル領域グループ、つまり波数グループを作製する。SFD で作製したスペクトル領域について Eq. (15) を用いて領域面積を算出する。その領域面積に *k*-means 法を適用することで、スペクトル領域をユークリッド距離に基づくスペクトル領域グループにクラスタリングする。SFD-*k*-means-PLS は各波数グループに含まれる全ての波数を用いて検量モデルを構築し、その SECV が小さい順に波数グループを選択する。SFD-*k*-means-PLS における調節パラメータは  $G_{\text{gen}}$  及び  $G_{\text{sel}}$  である。これらの調節パラメータは他の波数選択法と同様の範囲で最適化した。 $[G_{\text{gen}}, G_{\text{sel}}]$  の組合せは SECV が最小となるように選択した。

### 2.3.9 SFD-NCSC-PLS (提案法)

本章では SFD-NCSC-PLS を新規手法として提案する。

Fig. 11 に示すように、提案する SFD-NCSC 法は SFD 及び NCSC を用いてスペクトル領域グループを作製する。SFD は各吸収ピークに対応するスペクトル領域を作製することを意図している。その SFD に基づくスペクトル領域をその領域面積の相関関係に基づいてクラスタリングすることで、ある変動因子に由来する複数の吸収ピークに対応するスペクトル領域グループを作製できると考えられる。領域面積は SFD-*k*-means 法と同様に SFD を用いて算出する。NCSC を用いて、スペクトル領域を領域面積の相関に基づくスペクトル領域グループにクラスタリングする。各スペクトル領域グループに含まれる全ての波数を波数グループと定義する。

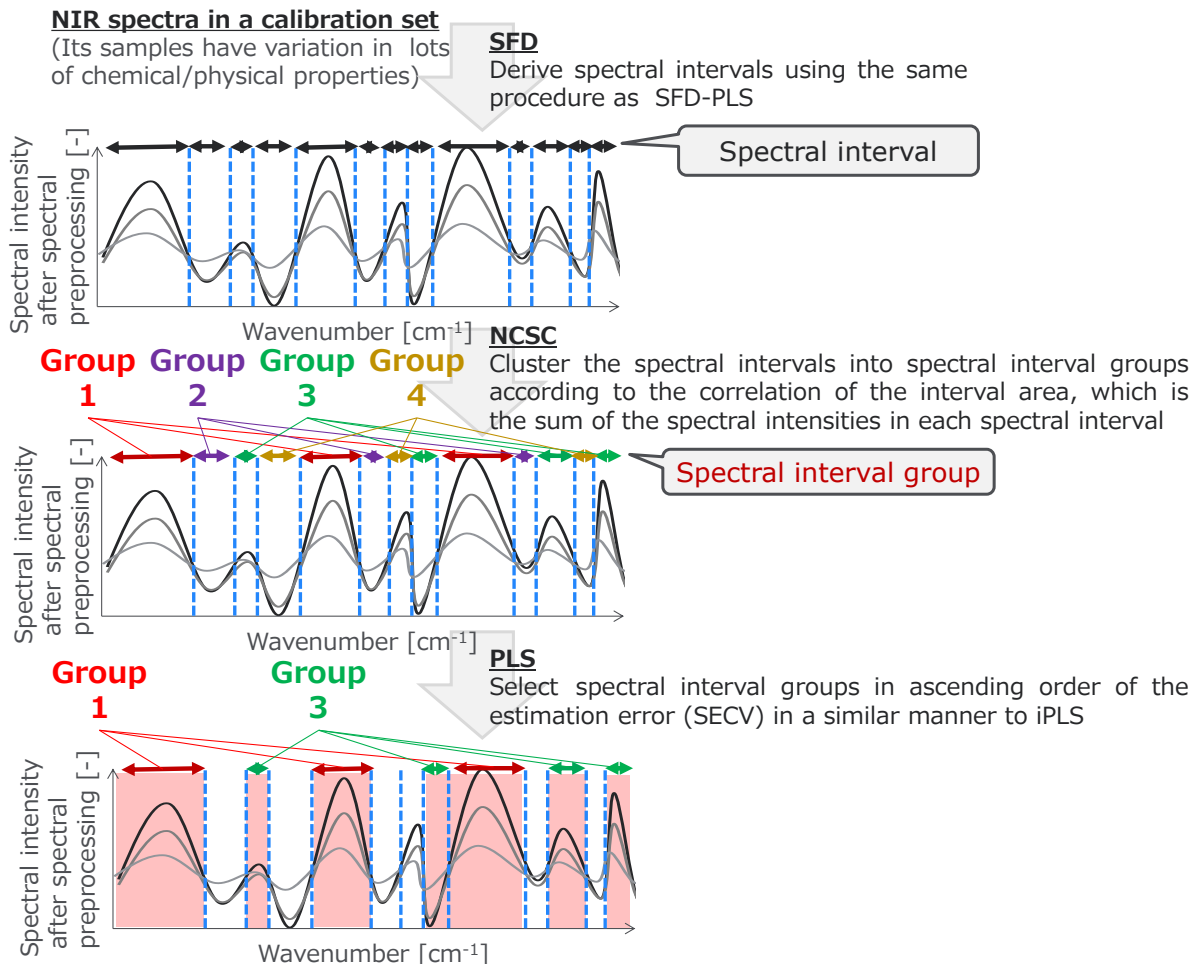


Fig. 11 An illustrative example of SFD-NCSC-PLS. [-] : Spectral intensity after spectral processing is dimensionless.

SFD-NCSC-PLS は、各波数グループに含まれる全ての波数を用いて検量モデルを構築し、その SECV が小さい順に波数グループを選択する。SFD-NCSC-PLS における調節パラメータは  $R_{\text{thres}}$ 、 $G_{\text{gen}}$ 、及び  $G_{\text{sel}}$  である。これらの調節パラメータは、他の波数選択法と同様の範囲で最適化した。SFD-NCSC-PLS に基づく波数選択手順を以下に示す:

(SFD の部分)

- 1) キャリブレーションセットにおける各波数の入力変数 (スペクトル強度) の標準偏差を算出し、スペクトル変動特性とする。
- 2) スペクトル全体をスペクトル変動特性の極小点で分割し、 $I_{\text{div}}$  個のスペクトル領域とする。

(NCSC の部分)

- 3) SFD で作製したスペクトル領域について Eq. (15) を用いて領域面積を算出する。 $R_{\text{thres}}=R_{\text{thres1}}$  及び  $G_t=G_{\text{min}}$  とする。ここで  $G_t$  は  $G_{\text{gen}}$  の暫定値である。
- 4) 領域面積に NC 法を適用することで  $S$  を算出する。
- 5)  $S$  に SC 法を適用することで  $G_t$  個の波数グループを暫定的に作製する。暫定的な



波数グループは波数を含まない可能性があるため、少なくとも1つの波数を含む暫定的な波数グループを  $G_{\text{gen}}$  個の波数グループと定義する。この時までにもし  $G_{\text{gen}}$  個の波数グループが既に作製されていれば、step 10 に進む。

(PLS の部分)

- 6) 各波数グループに含まれる全ての波数を用いて検量モデルを構築し、SECV を算出する。
- 7) 波数グループを SECV が小さい順に並べる。  $G_{\text{sel}}=1$  とする。
- 8) 最初の  $G_{\text{sel}}$  個の波数グループに含まれる全ての波数を用いて検量モデルを構築し、SECV を算出する。
- 9)  $G_{\text{sel}}=G_{\text{sel}}+1$  とし、  $G_{\text{sel}}=G_{\text{gen}}+1$  となるときまで step 8 に進む。(  $G_{\text{sel}}=G_{\text{gen}}$  である場合は PLS-All に相当するので省略する)
- 10)  $G_t=G_t+1$  とし、  $G_t=G_{\text{max}}$  となるときまで step 5 に進む。
- 11) もし  $R_{\text{thres}}=R_{\text{thres}1}$  であれば、  $R_{\text{thres}}=R_{\text{thres}2}$  及び  $G_t=G_{\text{min}}$  とし、step 4 に進む。
- 12) SECV が最小となる  $[R_{\text{thres}}, G_{\text{gen}}, G_{\text{sel}}]$  の最適な組合せを選択し、最初の  $G_{\text{sel}}$  個の波数グループに含まれる全ての波数を採用する。

## 2.4 結果及び考察

### 2.4.1 波数選択及び検量モデル構築

本項では、「2.3 波数選択法」項に示した9つの波数選択法を顆粒中の水分含量及び薬物(化合物 X)含量を推定する検量モデルの開発に適用し、それらの波数選択法を推定精度の観点で比較する。

#### 2.4.3.1 水分含量推定

選択された調節パラメータ及び推定精度を Table 8 に示し、選択された波数を Fig. 12 に図示する。キャリブレーションセットに関しては「1.4.3.1 水分含量推定」に記載の通りであるが、以下に改めて概説する。水分は  $11800\text{ cm}^{-1}$ 、 $10200\text{ cm}^{-1}$ 、 $8600\text{ cm}^{-1}$ 、 $7000\text{ cm}^{-1}$ 、 $5600\text{ cm}^{-1}$ 、及び  $5200\text{ cm}^{-1}$  付近に NIR 吸収帯を有する<sup>43)</sup>。Fig. 12 (b, c) に示すように、これらの吸収帯のうち  $7000\text{ cm}^{-1}$  及び  $5200\text{ cm}^{-1}$  付近の吸収帯は、キャリブレーションセットにおける  $7150\text{ cm}^{-1}$  及び  $5300\text{ cm}^{-1}$  付近の主要な吸収ピークと一致している。従って、これらの主要な吸収ピークは自由水に由来すると考えられた。キャリブレーションセットにおける  $5000\text{ cm}^{-1}$  付近の双峰性ピークも水分と関連付けることができるが、この双峰性ピークは自由水の影響だけではなく原薬(水和物である化合物 X)における結合水の影響をも受けている。一般に、水分に関連する波数領域におけるスペクトル強度は水分含量だけでなく薬物(化合物 X)含量や顆粒の物理的特性等の水分とは無関係な特性からの影響を受ける。水分と無関係な特性の変動が推定精度に与える影響を最小化するために、水分に関連する波数領域だけでなく水分と無関係の波数領域を選択することが水分含量を推定する検量モデル構築に有用で

ある<sup>12)</sup>。

iPLS は NIR スペクトル全体を 38 個のスペクトル領域に分割し、そのうち 5 個のスペクトル領域を選択した。SCMWPLS は NIR スペクトル全体を 19 個のスペクトル領域に分割し、それらのスペクトル領域に含まれる波数点のうち 308 波数点を選択した。SFD-PLS は NIR スペクトル全体を 207 個のスペクトル領域に分割し、そのうち 57 個のスペクトル領域を選択した。k-means-PLS は 2201 個の波数点 (NIR スペクトル全体) を 4 個の波数グループにクラスタリングし、そのうち 2 個の波数グループを選択した。NCSC-PLS は 2201 個の波数点 (NIR スペクトル全体) を 6 個の波数グループにクラスタリングし、そのうち 2 個の波数グループを選択した。interval-k-means-PLS は、まず NIR スペクトル全体を 7 個のスペクトル領域に分割し、次にそれらのスペクトル領域を 5 個のスペクトル領域グループにクラスタリングし、そしてそのうち 3 個のスペクトル領域グループを選択した。interval-NCSC-PLS は、まず NIR スペクトル全体を 17 個のスペクトル領域に分割し、次にそれらのスペクトル領域を 3 個のスペクトル領域グループにクラスタリングし、そしてそのうち 1 個のスペクトル領域グループを選択した。SFD-k-means-PLS は、まず NIR スペクトル全体を 207 個のスペクトル領域に分割し、次にそれらのスペクトル領域を 5 個のスペクトル領域グループにクラスタリングし、そしてそのうち 1 個のスペクトル領域グループを選択した。SFD-NCSC-PLS は、まず NIR スペクトル全体を 207 個のスペクトル領域に分割し、次にそれらのスペクトル領域を 7 個のスペクトル領域グループにクラスタリングし、そしてそのうち 2 個のスペクトル領域グループを選択した。

スペクトル領域に基づく波数選択法である iPLS、SCMWPLS、及び SFD-PLS は、 $7150\text{ cm}^{-1}$  及び  $5300\text{ cm}^{-1}$  付近の水分関連領域を含むスペクトル領域を作製したことから、これらの手法は水分含量の変動を反映した適切なスペクトル領域を作製できたと考えられる。さらに、これらの手法は水分関連領域だけでなく水分と関連しない領域も選択した。SFD-PLS 及び SCMWPLS は iPLS よりも SEP を低減したが、これはおそらく SFD-PLS 及び SCMWPLS が水分関連領域及び水分と関連しないが有用な領域を iPLS よりも適切に選択したためである。

波数クラスタリングに基づく波数選択法である k-means-PLS 及び NCSC-PLS は、 $12500\text{ cm}^{-1}$  から  $4200\text{ cm}^{-1}$  の 1 つの幅広いスペクトル領域及び  $4000\text{ cm}^{-1}$  付近の幾つかのスペクトル領域からなる波数グループを作製した。この幅広いスペクトル領域が生成された原因は、 $4000\text{ cm}^{-1}$  付近の波数領域における顕著な測定ノイズに由来する。Fig. 12 (a) に示すように、この測定ノイズにより  $4000\text{ cm}^{-1}$  付近の波数領域におけるスペクトル強度が極度に大きくなり、これによって  $12500\text{ cm}^{-1}$  から  $4200\text{ cm}^{-1}$  の波数領域におけるスペクトル強度が相対的に小さくなった。この相対的に小さいスペクトル強度はゼロ付近のほぼ一定値とみなすことができるため、対応する波数は全て同一の波数グループにクラスタリングされた。その結果として、幅広いスペクトル領域が生成された。

この幅広いスペクトル領域が生成された原因を考察するために、 $k$ 、 $l$ 、及び  $m$  の 3 つの波数点を考えることとする。ここで  $k$  及び  $l$  を  $12500\text{ cm}^{-1}$  から  $4200\text{ cm}^{-1}$  のスペクトル領域における波数位置、 $m$  を  $4000\text{ cm}^{-1}$  付近の波数位置とする。このとき、 $\|\mathbf{x}_k\| \ll \|\mathbf{x}_m\|$  及び  $\|\mathbf{x}_l\| \ll \|\mathbf{x}_m\|$  である。 $k$ -means 法では、 $\|\mathbf{x}_k - \mathbf{x}_l\| \leq \|\mathbf{x}_k\| + \|\mathbf{x}_l\| \ll \|\mathbf{x}_m\| \leq \|\mathbf{x}_k - \mathbf{x}_m\|$  であるため、 $k$  番目の波数と  $l$  番目の波数との間のスペクトル強度のユークリッド距離は  $k$  番目または  $l$  番目のいずれかの波数と  $m$  番目の波数との間のスペクトル強度のユークリッド距離よりも著しく小さい。従って、 $k$  番目及び  $l$  番目の波数は常に同じ波数グループにクラスタリングされる。NCSC では、 $[\mathbf{x}_{mk}, \mathbf{x}_{ml}] = [\mathbf{x}_k - \mathbf{x}_m, \mathbf{x}_l - \mathbf{x}_m] = [-\mathbf{x}_m, -\mathbf{x}_m]$  であるため、常に  $|C_{k,l}| = 1$  となり、 $k$  番目及び  $l$  番目の波数間の類似度である  $S(k, l)$  が大きくなる。従って、 $k$  番目及び  $l$  番目の波数は常に同じ波数グループにクラスタリングされる。結果として、 $k$ -means 法及び NCSC の両方において  $12500\text{ cm}^{-1}$  から  $4200\text{ cm}^{-1}$  の領域における全ての波数は同一の波数グループにクラスタリングされた。Fig. 12 (d) に示すように、波数クラスタリングに基づく 2 つの波数選択法は PLS-All とほぼ同様の波数を選択した。また、Table 8 に示すように、推定精度も PLS-All と比較して改善されなかったため、 $k$ -means-PLS 及び NCSC-PLS は実質的に適切な波数を選択することが出来なかった。これらの波数クラスタリングに基づく波数選択法を活用するためには、水分に関連する波数領域及び水分に関連しないが有用な波数領域を適切に選択できるようにクラスタリングのアルゴリズムや前処理法を改善する必要がある。

領域クラスタリングに基づく波数選択法である interval- $k$ -means-PLS、interval-NCSC-PLS、SFD- $k$ -means-PLS、及び SFD-NCSC-PLS において、スペクトル領域に基づく波数選択法と同様に、生成された  $7150\text{ cm}^{-1}$  及び  $5300\text{ cm}^{-1}$  付近におけるスペクトル領域グループは水分含量を反映していた。本章で評価した全ての手法の中で、SFD-NCSC-PLS は SEP 及び SEP と SEC との残差を最小化した。これはおそらく SFD-NCSC-PLS が水分に関連する波数領域及び水分に関連しないが推定精度向上に有用な波数領域を最も適切に選択したためである。以上から、SFD-NCSC-PLS は推定精度を最も向上させるとともにキャリブレーションセットにオーバーフィッティングする危険性を最も低減させた。これらの結果から、SFD-NCSC-PLS は SFD に基づくスペクトル領域を推定精度の向上に有用なスペクトル領域グループにクラスタリングできたことが示唆された。

Table 8 Comparison of wavenumber selection methods in the water content estimation (spectral preprocessing method: first derivative).

	$R_{\text{thres}}$	$G_{\text{sel}}/G_{\text{gen}}/I_{\text{div}}$	$M$	$K$	SECV [%]	SEC [%]	SEP [%]	$R^2$
PLS–All	-	-	2201	18	0.30	0.16	0.36	0.98
iPLS	-	5/38/38	285	4	0.24	0.22	0.32	0.99
SCMWPLS	-	-/-/19	308	6	0.24	0.19	0.29	0.99
SFD–PLS	-	57/207/207	978	6	0.25	0.19	0.29	0.99
$k$ -means–PLS	-	2/4/-	2168	13	0.27	0.17	0.37	0.98
NCSC–PLS	0.99	2/6/-	2168	13	0.27	0.17	0.37	0.98
interval– $k$ -means–PLS	-	3/5/7	1570	4	0.24	0.19	0.33	0.98
interval–NCSC–PLS	0.90	1/3/17	258	9	0.24	0.17	0.38	0.98
SFD– $k$ -means-PLS	-	1/5/207	79	4	0.26	0.24	0.39	0.98
SFD–NCSC–PLS	0.90	2/7/207	623	5	0.26	0.18	0.28	0.99

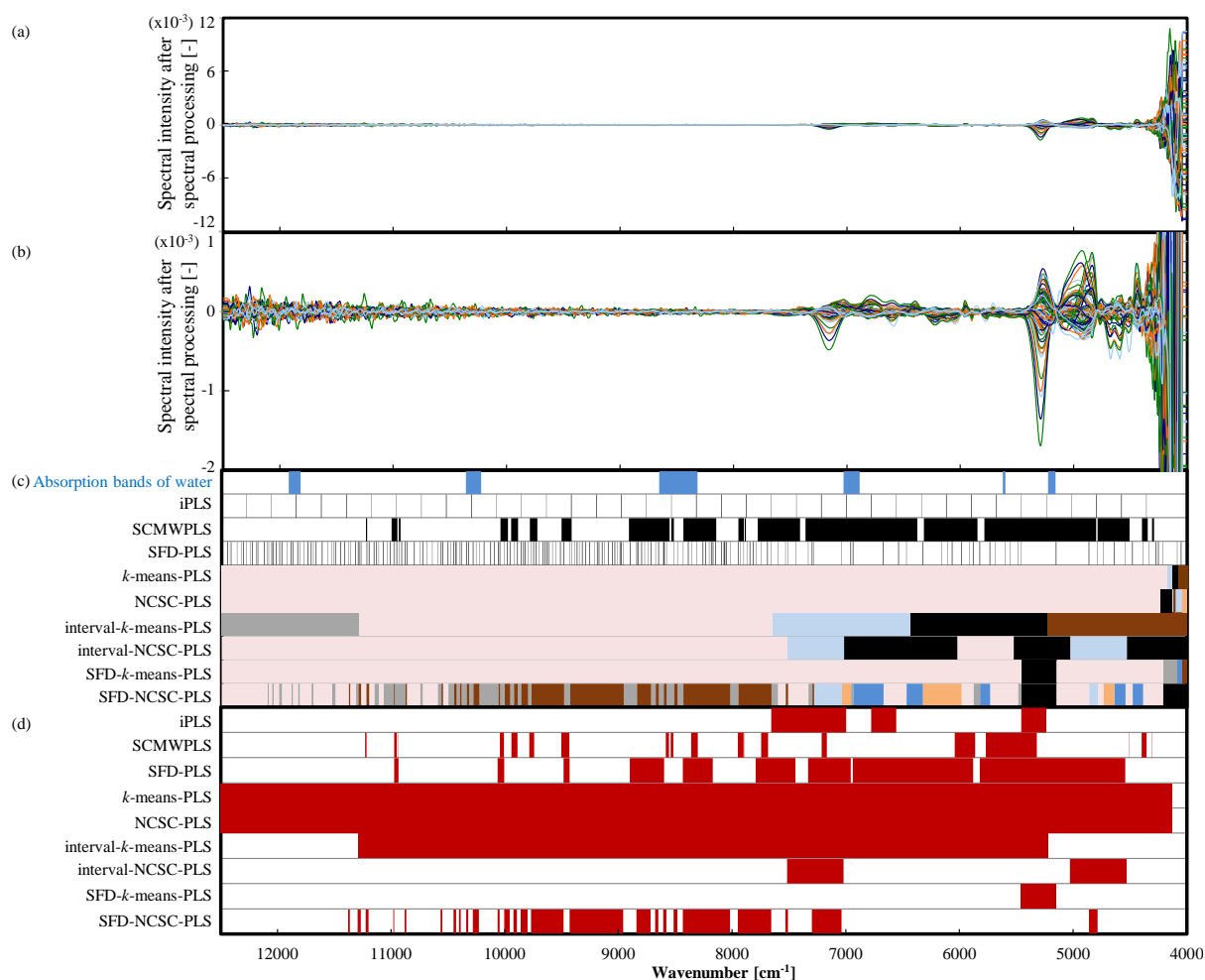


Fig. 12 Wavenumber clustering and selection results in the water content estimation (spectral preprocessing method: first derivative). (a) the preprocessed NIR spectra in the calibration set. (b) enlarged view of the preprocessed NIR spectra in the calibration set. (c) absorption bands of water and the generated spectral intervals and wavenumber groups; the same color means the same group in each method. In iPLS ( $I_{div}=38$ ) and SFD-PLS ( $I_{div}=207$ ), the verticle lines represent boundaries of spectral intervals. In SCMWPPLS ( $I_{div}=19$ ), the shaded regions represent spectral intervals. In  $k$ -means-PLS ( $G_{gen}=4$ ), NCSC-PLS ( $G_{gen}=6$ ), interval- $k$ -means-PLS ( $G_{gen}=5$ ), interval-NCSC-PLS( $G_{gen}=3$ ), SFD- $k$ -means-PLS ( $G_{gen}=5$ ), and SFD-NCSC-PLS ( $G_{gen}=7$ ), the shaded regions represent wavenumber groups. (d) the selected wavenumbers (shaded regions). [-] : Spectral intensity after spectral processing is dimensionless.

#### 2.4.3.2 薬物含量推定

選択された調節パラメータ及び推定精度を Table 9 に示し、選択された波数を Fig. 13 に図示する。キャリブレーションセットに関しては「1.4.3.2 薬物含量推定」に記載の通りであるが、以下に改めて概説する。Fig. 13 (b) に示すように、薬物（化合物 X）は  $6500\text{ cm}^{-1}$ 、 $5150\text{ cm}^{-1}$ 、 $4800\text{ cm}^{-1}$ 、 $4500\text{ cm}^{-1}$ 、及び  $4100\text{ cm}^{-1}$  付近において主要な吸収ピークを示す。Fig. 13 (a, b) に示すように、これらの薬物（化合物 X）の吸収ピークのうち  $6500\text{ cm}^{-1}$  及び  $5150\text{ cm}^{-1}$  付近の吸収ピークは、キャリブレーション

ンセットにおける主要な吸収ピークと一致している。7150  $\text{cm}^{-1}$ 、5300  $\text{cm}^{-1}$ 、及び 5000  $\text{cm}^{-1}$  付近の波数領域において顕著な変動を示す NIR スペクトルは、高い水分含量を有するキャリブレーション用のサンプルに相当する。従って、「2.4.3.1 水分含量推定」項に記載の通り、7150  $\text{cm}^{-1}$  及び 5300  $\text{cm}^{-1}$  付近のスペクトル強度は自由水を反映し、5000  $\text{cm}^{-1}$  付近のスペクトル強度は自由水及び薬物（化合物 X）における結合水を反映していると考えられる。水分含量の変動を予め検量モデルに組み込み、水分含量の変動が推定精度に与える影響を低減させる目的で、この高水分含量のサンプルを意図的にキャリブレーションに採用した。水分含量の場合と同様に、薬物（化合物 X）に関連する波数領域におけるスペクトル強度は薬物（化合物 X）含量だけでなく水分含量や顆粒の物理的特性等の薬物（化合物 X）とは無関係な特性の影響を受ける。薬物（化合物 X）と無関係な特性の変動が推定精度に与える影響を最小化するために、薬物（化合物 X）に関連する波数領域だけでなく薬物（化合物 X）と無関係な波数領域を選択することが薬物（化合物 X）含量を推定する検量モデル構築に有用であると考えられた。

iPLS は NIR スペクトル全体を 38 個のスペクトル領域に分割し、そのうち 27 個のスペクトル領域を選択した。SCMWPLS は NIR スペクトル全体を 17 個のスペクトル領域に分割し、それらのスペクトル領域に含まれる波数点のうち 186 波数点を選択した。SFD-PLS は NIR スペクトル全体を 156 個のスペクトル領域に分割し、そのうち 28 個のスペクトル領域を選択した。k-means-PLS は 2202 個の波数点（NIR スペクトル全体）を 6 個の波数グループにクラスタリングし、そのうち 4 個の波数グループを選択した。NCSC-PLS は 2202 個の波数点（NIR スペクトル全体）を 4 個の波数グループにクラスタリングし、そのうち 1 個の波数グループを選択した。interval-k-means-PLS は、まず NIR スペクトル全体を 15 個のスペクトル領域に分割し、次にそれらのスペクトル領域を 7 個のスペクトル領域グループにクラスタリングし、そしてそのうち 2 個のスペクトル領域グループを選択した。interval-NCSC-PLS は、まず NIR スペクトル全体を 51 個のスペクトル領域に分割し、次にそれらのスペクトル領域を 7 個のスペクトル領域グループにクラスタリングし、そしてそのうち 2 個のスペクトル領域グループを選択した。SFD-k-means-PLS は、まず NIR スペクトル全体を 156 個のスペクトル領域に分割し、次にそれらのスペクトル領域を 5 個のスペクトル領域グループにクラスタリングし、そしてそのうち 2 個のスペクトル領域グループを選択した。SFD-NCSC-PLS は、まず NIR スペクトル全体を 156 個のスペクトル領域に分割し、次にそれらのスペクトル領域を 6 個のスペクトル領域グループにクラスタリングし、そしてそのうち 3 個のスペクトル領域グループを選択した。

スペクトル領域に基づく波数選択法である iPLS、SCMWPLS、及び SFD-PLS は 6500  $\text{cm}^{-1}$  及び 5150  $\text{cm}^{-1}$  付近における薬物（化合物 X）の主要な吸収ピークを含むスペクトル領域を作製した。従って、これらの手法は薬物（化合物 X）含量の変動を反映する適切なスペクトル領域を作製したと考えられる。スペクトル領域に基づく波数

選択法の中で、SFD-PLS は SEP を最小化させた。さらに SFD-PLS は水分含量の変動に対する頑健性を向上するように  $7150\text{ cm}^{-1}$ 、 $5300\text{ cm}^{-1}$ 、及び  $5000\text{ cm}^{-1}$  付近の自由水及び結合水が関連する波数領域を除外した。これらの結果から、SFD-PLS は iPLS 及び SCMWPLS よりも正確に薬物(化合物 X)に関連する波数領域及び薬物(化合物 X)と無関係であるが推定精度の向上に有用な波数領域の両方を選択したと推察される。

波数クラスタリングに基づく波数選択法である *k*-means-PLS 及び NCSC-PLS において、スペクトル領域に基づく波数選択法と同様に、 $6500\text{ cm}^{-1}$  及び  $5150\text{ cm}^{-1}$  付近で生成された波数グループは薬物(化合物 X) 含量の変動を反映したと考えられる。NCSC-PLS は  $6500\text{ cm}^{-1}$  及び  $5150\text{ cm}^{-1}$  付近の薬物(化合物 X) に関連する波数領域を含む波数グループと、 $7150\text{ cm}^{-1}$  及び  $5300\text{ cm}^{-1}$  付近の自由水に関連する波数領域を含む別の波数グループを生成した。従って、NCSC-PLS は薬物(化合物 X) に関連する波数領域と自由水に関連する波数領域を別々に反映する波数グループを生成した。波数グループを選択する際、NCSC-PLS は SFD-PLS と同様に自由水及び結合水に関連する波数領域を除外した。NCSC-PLS は *k*-means-PLS よりも SEP を低減させたが、これはおそらく NCSC-PLS が *k*-means-PLS よりも正確に薬物(化合物 X) 含量に関連する波数領域及び薬物(化合物 X) と無関係であるが推定精度の向上に有用な波数領域の両方を選択したためである。

領域クラスタリングに基づく波数選択法である interval-*k*-means-PLS、interval-NCSC-PLS、SFD-*k*-means-PLS、及び SFD-NCSC-PLS において、 $6500\text{ cm}^{-1}$  及び  $5150\text{ cm}^{-1}$  付近で生成されたスペクトル領域グループは薬物(化合物 X) 含量の変動を反映したと考えられる。この結果はスペクトル領域に基づく波数選択法及び波数クラスタリングに基づく波数選択法と同様である。SFD-NCSC-PLS は、NCSC-PLS と同様に、薬物(化合物 X) 及び自由水に関連する波数領域を別々に反映する2つのスペクトル領域グループを生成した。スペクトル領域グループを選択する際、SFD-NCSC-PLS は自由水及び結合水に関連する  $7150\text{ cm}^{-1}$ 、 $5300\text{ cm}^{-1}$ 、及び  $5000\text{ cm}^{-1}$  付近の波数領域を除外したが、この結果は SFD-PLS 及び NCSC-PLS と同様である。

水分含量推定における結果と一致して、本章で評価した全ての手法の中で、SFD-NCSC-PLS は SEP 及び SEP と SEC との残差の両方を最小化した。従って、SFD-NCSC-PLS は薬物(化合物 X) に関連する波数領域及び薬物(化合物 X) と無関係であるが推定精度の向上に有用な波数領域の両方を最も正確に選択したと考えられる。これらの結果から、SFD-NCSC-PLS は SFD に基づくスペクトル領域を推定精度の向上に有用なスペクトル領域グループにクラスタリングできたことが示唆される。

水分含量及び薬物(化合物 X) 含量を推定する2つの事例における結果は、提案する SFD-NCSC-PLS が推定精度の向上及びキャリブレーションセットにオーバーフィッティングする危険性の両観点から最も優れていることを示している。

Table 9 Comparison of wavenumber selection methods in the drug content estimation (spectral preprocessing method: first derivative + SNV).

	$R_{\text{thres}}$	$G_{\text{sel}}/G_{\text{gen}}/I_{\text{div}}$	$M$	$K$	SECV [%]	SEC [%]	SEP [%]	$R^2$
PLS–All	-	-	2202	20	2.42	0.13	2.93	0.88
iPLS	-	27/38/38	1575	20	1.45	0.25	2.53	0.92
SCMWPLS	-	-/-/17	186	8	2.54	1.68	2.18	0.94
SFD–PLS	-	28/156/156	649	13	1.57	1.03	1.81	0.96
$k$ -means–PLS	-	4/6/-	459	13	1.62	1.06	2.25	0.93
NCSC–PLS	0.90	1/4/-	303	13	1.58	1.07	2.00	0.95
interval– $k$ -means–PLS	-	2/7/15	742	19	1.31	0.37	2.43	0.92
interval–NCSC–PLS	0.90	2/7/51	817	20	1.37	0.17	2.94	0.90
SFD– $k$ -means–PLS	-	2/5/156	298	20	1.71	0.55	3.34	0.86
SFD–NCSC–PLS	0.99	3/6/156	857	13	1.52	0.99	1.69	0.96



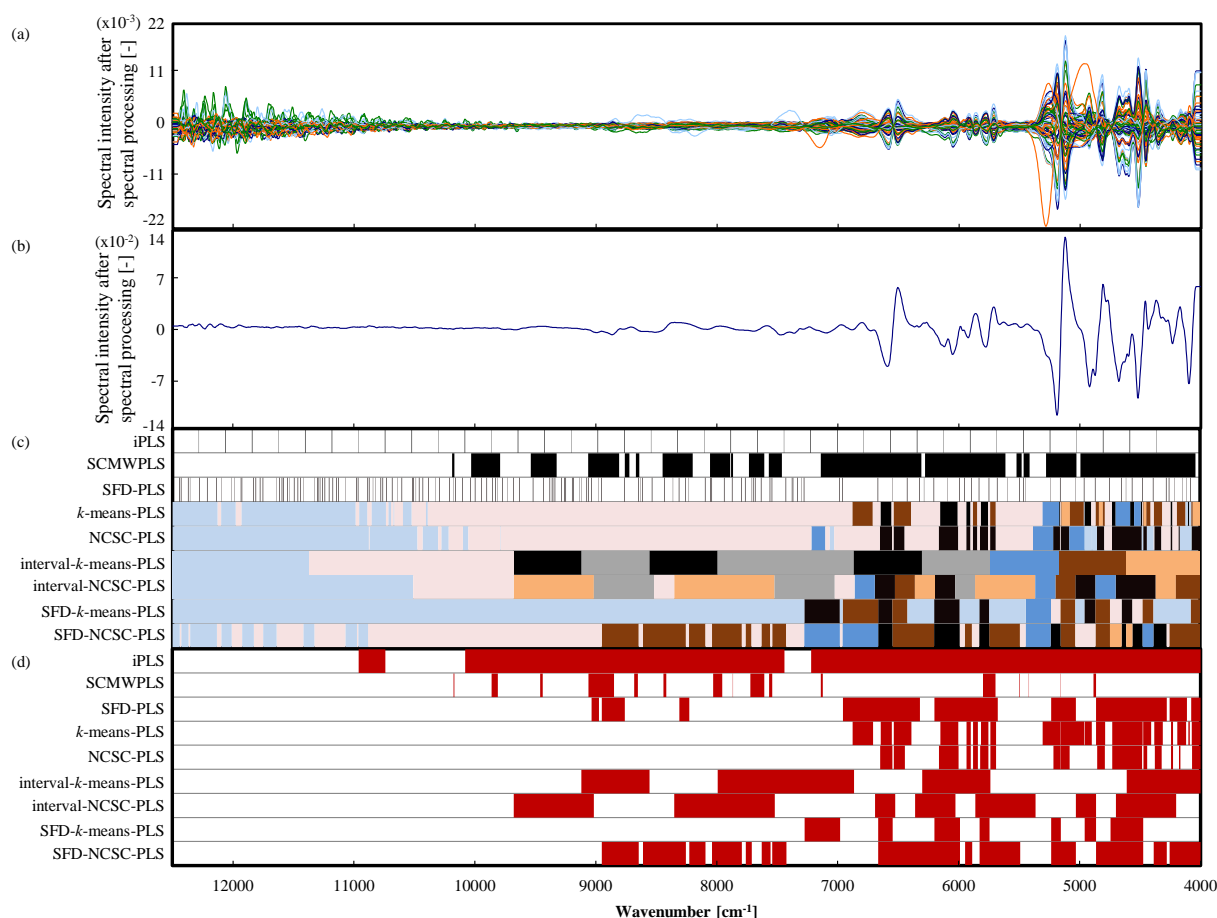


Fig. 13 Wavenumber clustering and selection results in the drug content estimation (spectral preprocessing method: first derivative + SNV). (a) the preprocessed NIR spectra in the calibration set. (b) the preprocessed NIR spectrum of the drug substance. (c) the generated spectral intervals and wavenumber groups; the same color means the same group in each method. In iPLS ( $I_{\text{div}}=38$ ) and SFD-PLS ( $I_{\text{div}}=156$ ), the verticle lines represent boundaries of spectral intervals. In SCMWPPLS ( $I_{\text{div}}=17$ ), the shaded regions represent spectral intervals. In  $k$ -means-PLS ( $G_{\text{gen}}=6$ ), NCSC-PLS ( $G_{\text{gen}}=4$ ), interval- $k$ -means-PLS ( $G_{\text{gen}}=7$ ), interval-NCSC-PLS ( $G_{\text{gen}}=7$ ), SFD- $k$ -means-PLS ( $G_{\text{gen}}=5$ ), and SFD-NCSC-PLS ( $G_{\text{gen}}=6$ ), the shaded regions represent wavenumber groups. (d) the selected wavenumbers (shaded regions). [-] : Spectral intensity after spectral processing is dimensionless.

## 2.4.2 計算時間

Table 10 及び Table 11 に示すように、計算時間は水分含量及び薬物（化合物 X）含量を推定する両事例について同様の傾向を示した。総計算時間は、スペクトル領域の作製に要した時間、 $k$ -means 法または NCSC に基づくクラスタリングに要した時間、及び PLS モデル化の繰り返しに要した時間に細分化できる。評価する波数の組合せの数が多くなるほど、そして PLS モデルを一回構築するために要する計算時間が大きくなるほど、PLS モデル化の繰り返しに長い計算時間を要する。また、入力変数の数（使用する波数点の数）が増えるほど、PLS モデルを一回構築するために要する時間は指数関数的に長くなる。

スペクトル領域に基づく波数選択法では、PLS モデル化の繰り返しの要した時間が総計算時間の 99%以上を占めており、スペクトル領域の作製に要した時間はごく僅かであった。SCMWPLS は、選択する波数位置の数を  $N_p=M_{all}/2$  に固定したため、各 PLS モデル化に使用した波数点の数は iPLS 及び SFD-PLS の 1/10 以下であった。この結果、SCMWPLS において PLS モデルを一回構築するために要する計算時間は著しく短くなった。従って、SCMWPLS は最も多い回数の波数の組合せ評価を実施したにも関わらず、PLS モデル化の繰り返し構築に要する時間は最も短くなった。iPLS 及び SFD-PLS においては、各 PLS モデルの構築に同程度の数の波数点を用いられた。iPLS は網羅的な範囲の  $I_{div}$  を評価したため、4949 回の波数の組合せ評価が必要となった。一方、SFD-PLS は  $I_{div}$  を一意的かつ自動的に決定するため、約 200 回の波数の組合せ評価で波数選択が完了した。結果として、SFD-PLS における PLS モデル化の繰り返しの要した計算時間は iPLS よりも著しく短くなった。ここで、もし生成されたスペクトル領域をスペクトル領域グループにクラスタリングすれば、波数の組合せの数が低減するため、この PLS モデル化の繰り返しの要する計算時間をさらに短くできるであろう。この方策は後述の領域クラスタリングに基づく波数選択法で検討する。

波数クラスタリングに基づく波数選択において、NCSC は  $k$ -means 法よりも著しく長い計算時間を要した。NCSC-PLS は  $G_{gen}$  及び  $G_{sel}$  の各組合せについて複数の  $R_{thres}$  を評価したため、NCSC-PLS はより多い回数の波数の組合せ評価を必要とした。さらに NCSC は、膨大な数の変数をクラスタリングする場合、Eq. (18) を解くために非常に大きな計算負荷がかかる。ここで、もしスペクトル全体をスペクトル領域に分割し、個々の波数の代わりにスペクトル領域を用いてクラスタリングすれば、クラスタリングする変数の数が低減するため、NCSC の計算時間を低減できるであろう。この方策は後述の領域クラスタリングに基づく波数選択法で検討する。

領域クラスタリングに基づく波数選択法では、PLS モデル化の繰り返しの要した時間が総計算時間の 99%以上を占めており、スペクトル領域の作製及び  $k$ -means 法または NCSC に基づくクラスタリングに要した時間はごく僅かであった。SFD-NCSC-PLS は SFD-PLS より少ない数の波数の組合せを評価したため、PLS モデル化の繰り返しの要する時間を半分以下に短縮した。さらに、SFD-NCSC-PLS は NCSC-PLS と比較してクラスタリングする変数の数を減少させたため、NCSC に基づくクラスタリングに要する時間を 1/15 以下に短縮させた。結果として、SFD-NCSC-PLS は PLS モデル化の繰り返し及び NCSC に基づくクラスタリングに必要な計算負荷を大幅に低減した。

以上をまとめると、SFD-NCSC-PLS は適切に波数領域を選択することで最も高い推定精度を達成すると同時に、計算負荷を低減させた。従って、SFD-NCSC-PLS は推定精度及び計算負荷の両観点から本検討で評価した波数選択法の中で最良の手法である。

Table 10 Computation time, the number of wavenumbers used in each PLS modeling (*M/PLS*), and the number of evaluated combinations of wavenumbers (#eval) in the water content estimation.

	Making intervals <sup>a</sup>	Computation time [h]				Total	<i>M/PLS</i>	#eval
		<i>k</i> -means	NCSC	PLS	(Average [Min, Max])			
iPLS	<0.01	-	-	56.13	56.13	1084 [22, 2179]	4949	
SCMWPLS	<0.01	-	-	0.92	0.92	72 [1, 308]	88406	
SFD-PLS	<0.01	-	-	1.24	1.24	1302 [63, 2192]	206	
<i>k</i> -means-PLS	-	0.01	-	0.72	0.73	2172 [2152, 2193]	20	
NCSC-PLS	-	-	0.48	0.84	1.32	2013 [2, 2195]	27	
interval- <i>k</i> -means-PLS	<0.01	0.01	-	40.08	40.09	1440 [25, 2158]	1910	
interval-NCSC-PLS	<0.01	-	0.04	45.62	45.66	1195 [24, 2170]	3164	
SFD- <i>k</i> -means-PLS	<0.01	<0.01	-	0.62	0.62	1845 [79, 2192]	20	
SFD-NCSC-PLS	<0.01	-	<0.01	0.59	0.59	1360 [83, 2127]	34	

a : Computation time to build PLS models was excluded.

Table 11 Computation time, the number of wavenumbers used in each PLS modeling (*M/PLS*), and the number of evaluated combinations of wavenumbers (#eval) in the drug content estimation.

	Making intervals <sup>a</sup>	Computation time [h]				Total	<i>M/PLS</i>	#eval
		<i>k</i> -means	NCSC	PLS	(Average [Min, Max])			
iPLS	<0.01	-	-	37.61	37.61	1113 [22, 2180]	4949	
SCMWPLS	<0.01	-	-	0.50	0.50	66 [1, 186]	80621	
SFD-PLS	<0.01	-	-	1.46	1.46	1277 [26, 2198]	155	
<i>k</i> -means-PLS	-	<0.01	-	0.20	0.20	1233 [131, 2105]	20	
NCSC-PLS	-	-	0.47	0.57	1.04	1589 [181, 2185]	40	
interval- <i>k</i> -means-PLS	<0.01	0.01	-	21.43	21.44	1411 [37, 2179]	1910	
interval-NCSC-PLS	<0.01	-	0.05	29.69	29.74	1160 [24, 2150]	3549	
SFD- <i>k</i> -means-PLS	<0.01	<0.01	-	0.27	0.27	1410 [214, 2149]	20	
SFD-NCSC-PLS	<0.01	-	<0.01	0.39	0.39	1230 [188, 1996]	40	

a : Computation time to build PLS models was excluded.

## 2.5 小括

本章では、NIRS に基づく検量モデルを効率的に開発する技術として SFD 及び

NCSC を組み合わせた波数選択法を提案した。第一章では、小さな計算負荷で推定精度を改善する波数選択法として、スペクトル領域に基づく手法 **SFD-PLS** を提案した。**SFD-PLS** は各吸収ピークに対応するスペクトル領域を作製することを意図したものであり、従来法である **iPLS** よりも小さな計算負荷で高い推定精度を達成した。しかし、そのスペクトル領域の数は膨大であり、より効率的に高い推定精度を得る観点で改善の余地があった。そこで第二章では、各変動因子に由来する複数の吸収ピークに対応するスペクトル領域グループを作製することを意図した波数選択法 **SFD-NCSC-PLS** を新規に開発した。**SFD-NCSC-PLS** は、まず **SFD** によりスペクトル領域を作製し、そのスペクトル領域を **NCSC** により数個のスペクトル領域グループにクラスタリングする。そして、各スペクトル領域グループに含まれる全ての波数を用いて検量モデルを構築し、その検量モデルの推定精度が高い順にスペクトル領域グループを選択する。

事例研究として、提案する波数選択法を用いて顆粒中の水分含量及び薬物（化合物 **X**）含量を推定する 2 種類の検量モデルを構築し、推定精度及び計算時間の観点で従来の波数選択法と比較評価した。両事例において、**SFD-NCSC-PLS** は従来法である **iPLS** 及び **SCMWPLS** 並びに **SFD** または **NCSC** の一方のみを用いた手法よりも高い推定精度を達成した。この結果は、**SFD-NCSC-PLS** が推定対象とする品質特性（水分含量及び薬物含量）に関連する波数を最も正確に選択したことを示唆している。品質特性を推定する上で有用なスペクトル領域は様々な幅を有する。**iPLS** は等幅のスペクトル領域を作製するため、有用なスペクトル領域を柔軟に選択することができない。**SCMWPLS** は **MWPLS** 及び **CSMWPLS** を組合せた手法であり、作製するスペクトル領域の幅を調節可能である。しかし、**MWPLS** は移動枠を用いて各波数点を近似的に評価し、その評価結果に基づいてスペクトル領域を作製する。さらに **CSMWPLS** は **MWPLS** により作製した各スペクトル領域から唯一つの副スペクトル領域を抽出する。しかし **MWPLS** により作製した各スペクトル領域には推定精度を向上させるために有用なスペクトル領域が複数存在する可能性がある。こうした **SCMWPLS** のアルゴリズムでは、適切なスペクトル領域を選択し、推定精度を向上させることができない可能性がある。一方、**SFD-NCSC-PLS** は、各吸収ピークに対応するスペクトル領域を作製し、そのスペクトル領域をその領域面積の相関関係に基づいてクラスタリングすることで、ある変動因子に由来する複数の吸収ピークに対応するスペクトル領域グループを作製することを意図したものである。従って、**SFD-NCSC-PLS** は他の手法よりも正確に複数の吸収ピークに対応するスペクトル領域グループを選択することができたと考えられる。計算負荷は波数選択法を運用する上で実際的な課題の一つであるが、**SFD-NCSC-PLS** は **SFD** または **NCSC** の一方のみを用いた手法と比較して計算負荷を著しく低減させた。なお、**SFD-NCSC-PLS** は一般的な家庭用計算機を用いて 1 時間以内に波数選択を完了した。従って、**SFD-NCSC-PLS** は第一章で提案した **SFD-PLS** よりも小さな計算負荷で高い推定精度を達成するさらに有望な波数選

択法である。

今回、異なる2つの推定対象に対して同様の結果が得られたことから、広範な推定対象に対してこれらの提案法を適用できると考えられる。本検討では顆粒のみを測定対象としたが、一般に NIR 吸収帯に応じて吸収ピークが生じ、その吸収ピークが重なり合う等の NIR スペクトルの特徴は、測定対象に依らず同様である。従って、顆粒以外の様々な測定対象に対してもこれらの提案法を適用できるだろう。これらの提案法に基づく波数選択を実行するプログラムを作成しておけば、そのプログラムを実行することで自動的に波数選択を実施できる。従って、波数選択を実施する担当者に依らず同一の結果を得られる。また、推定対象や測定対象に依らず、同一のプログラムを用いて波数選択を実施できる。以上から、これらの提案法は推定精度及び計算負荷の観点で優れているだけでなく、波数選択結果が個人の能力に依存しない波数選択法であり、多様な推定対象及び測定対象に適用できる。

### 第三章 検量モデルメンテナンスの知識統合に基づく運用化

#### 3.1 序論

第三章では、NIRS に基づく検量モデルをメンテナンスする枠組み及びメンテナンスを知識統合に基づいて運用化することを提案する。

第一章及び第二章では、新規に提案した波数選択法を用いることで推定精度の高い検量モデルを構築できることを示した。しかし、検量モデルを開発した時点において検量モデルが所望の推定精度を有していたとしても、製品ライフサイクルの中でその推定精度が低下する可能性がある。これは NIR スペクトルが推定対象 (e.g., 薬物) 以外の化学物質 (e.g., 水分) に関する情報及び物理的な情報 (e.g., 測定物質の密度) を反映するためである。従って、開発段階では NIR スペクトルに影響を与える変動因子を予め検証し、推定対象以外の変動因子の影響を受けにくい頑健な検量モデルを構築する。しかし、製剤生産において原料や製造プロセスの特性が想定する範囲を超えて変動する可能性がある。例えば、製剤の原料である添加剤の供給業者を変更すると、その添加剤の物理的特性が変動し、それを用いて製造される製剤の物理的特性に影響を与え得る。このような外乱変動は製剤の NIR スペクトルに影響を与え、それによって検量モデルの推定結果に影響を与える可能性がある<sup>29)</sup>。そこで、推定精度を恒常的に保証するために製剤のライフサイクルを通じて検量モデルをメンテナンスすることが必要である。この検量モデルメンテナンスは ICH<sup>41)</sup>、EMA<sup>32)</sup>、欧州薬局方 (Ph. Eur.)<sup>42)</sup>、及び USP<sup>33)</sup>により要求されている。

検量モデルメンテナンスの重要性及び特定の評価手法が概念としては報告されているものの<sup>30), 31), 32)</sup>、その具体的な枠組みはこれまで報告されていない。もし各種のガイドライン<sup>31), 32)</sup>及び USP<sup>33)</sup>に準拠した検量モデルメンテナンスの具体的な枠組みが提案されれば、製剤を製造する商用生産工場にて検量モデルメンテナンスを運用する上での拠り所となるだろう。

その具体的な枠組みを踏まえた上で、検量モデルメンテナンスの運用化に向けては GMP を考慮する必要がある。検量モデルメンテナンスを含む製剤製造の業務は、GMP に従って運用される。GMP 管理下では、作業手順及び役割を適切に文書化し、その作業結果を遅滞なく記録することが求められる。従って、検量モデルメンテナンスを運用化するためには、詳細な作業手順及び役割を明確化することが重要である。GMP に従って運用する場合、製造管理、品質管理、品質保証、及び技術開発等の複数の部署が関与するため、各部署が有する運用に関する知識及び責任を十分考慮すべきである。もし各部署の運用に関する知識が適切に共有されていなければ、情報伝達が遅延する、運用方法に関して他部署と相談するための打合せが頻繁に発生するといった問題が発生する恐れがある。しかし、複数の部署が関与する新規の枠組みに対して、あらゆる事態に対応する詳細な作業手順及び役割を予め明確化することは困難である。また、GMP 管理下で新規の枠組みを運用化するための方法論が論じられることは極めて少ない。もしそのような方法論が提案されれば、GMP 管理下で検量モデ

ルメンテナンスを適切かつ効率的に運用化するために有用となるだろう。

詳細な作業手順及び役割を予め明確化するための方策の一つとして、各担当者の知識を共有しその知識を統合することが有望と考えられる。ここで知識を統合するとは、個々の担当者が有する知識を関連付け、文章や図表を用いた有形の知識にまとめることによって、あらゆる担当者が個々の担当者が有する知識を活用できる状態にすることである。検量モデルメンテナンスの運用に関して知識を統合することで、あらゆる部署の担当者がプロセス全体及び詳細な情報の流れを容易に理解することができ、運用を効率化できることが期待される。知識統合の手段として、業務プロセスモデル化手法である IDEF0 及び役割分担表 RACI を用いることができる。

IDEF0 は、業務及び情報の流れを可視化するための業務プロセスモデル化技術である<sup>49), 50)</sup>。製薬産業において IDEF0 を適用した事例として、QbD に基づく生物製剤を開発するための情報管理が報告されている<sup>51)</sup>。RACI は、各業務における役割及び責任を可視化するための役割分担表である<sup>52)</sup>。Sugiyama 及び Schmidt は、IDEF0 で定義した業務の役割及び責任を定義するために IDEF0 と RACI を相補的に組み合わせることを提案している<sup>53)</sup>。また、その適用事例として注射製剤製造工程における製品収率を改善したことを報告している。ただし、これらの業務プロセスモデル化手法は、検量モデルメンテナンスのような新規枠組みの運用化に向けた知識統合の手段として適用された事例は報告されていない。

本章では、医薬品生産において GMP に従って実施される検量モデルメンテナンスについて、その具体的な枠組み及び運用化に向けた知識統合に基づく方法論を新規に提案する。まず、規制ガイドライン<sup>31), 32)</sup>及び USP<sup>33)</sup>を参照しながら検量モデルメンテナンスの枠組みを作成した。次に、GMP を考慮して適切に検量モデルメンテナンスを運用するために、商用生産工場における各部署に所属する担当者が有する知識を統合した。その知識統合の手段として、業務プロセスモデル化技術である IDEF0 及び RACI を用いた。さらに、知識統合が検量モデルメンテナンスを運用する効率に与える効果を評価した。また、商用生産工場において検量モデルメンテナンスを実際に運用し、運用化した業務の流れが妥当であることを検証した。

## 3.2 実験の部

### 3.2.1 業務プロセスモデル化技術: IDEF0

IDEF0 は、作業及び情報の流れを可視化するための業務プロセスモデル化技術である<sup>49), 50)</sup>。IDEF0 モデルの概念図を Fig. 14 に示す。IDEF0 モデルは、長方形と矢印で構成される。長方形は作業内容を示し、その内容が動詞で始まる語句で記述される。矢印は情報の流れを示し、その方向に基づいて以下の四つに分類される：

- 1) 入力（左から）：その作業で使用する資源のうち、その作業中で出力に変換されるもの（例：現行の検量モデル）。
- 2) 制御（上から）：従わなければならない要件。本章では、その作業を開始す

る契機とする（例：モデル更新指示書）。

3) 機能（下から）：その作業で使用する資源のうち、その作業中に変化しないもの（例：標準操作手順書（Standard operating procedure: SOP））。

4) 出力（右へ）：その作業の成果物として得られるもの。その作業における出力の1つは作業の完了を意味する（例：更新後の検量モデル及び変更管理報告書）。

その作業を開始する契機及び完了時の成果物を確認する場合は、制御及び出力の矢印に注目すればよい。必要な資源を確認する場合は、入力及び機能の矢印に注目すればよい。Fig. 14 に示すように、作業をより詳細に記述するために、各作業を小作業に分割することができる。IDEF0 モデルを作図する上で、process industry executive for achieving business advantage using the standard for data exchange（PIEBASE）<sup>54)</sup>及び Sugiyama ら<sup>55)</sup>が提案する IDEF0 の定型書式を参考とした。IDEF0 モデルの作図には Microsoft Visio Professional 2010（Microsoft, Redmond, WA, USA）を用いた。

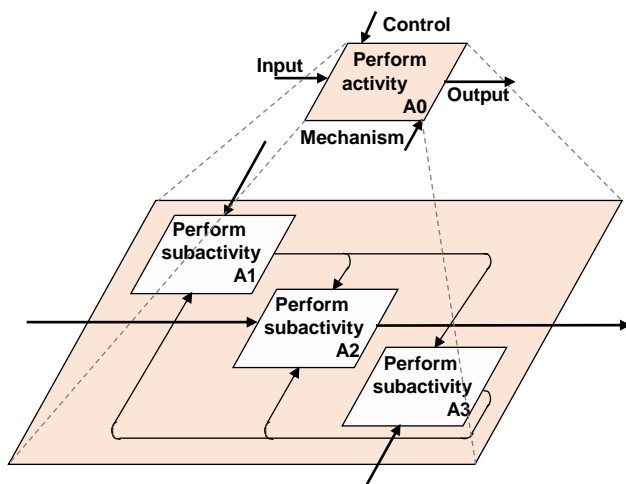


Fig. 14 Example of an IDEF0 model

### 3.2.2 役割分担表: RACI

IDEF0 は各作業における役割分担を定義しないため、IDEF0 と RACI とを相補的に組合せてその役割と責任を定義した。RACI は役割分担の記述手法であり、作業の役割及び責任を以下の四つに分類する：

- 1) R: 実行責任者。作業を遂行する者。主な役割は報告書の作成である。責任を明確化するために唯一人の担当者を割り当てる。
- 2) A: 説明責任者。作業を承認し、その作業に関する文書に署名する者。主な役割は報告書の承認である。責任を明確化するために唯一人の担当者を割り当てる。
- 3) C: 相談先。必要に応じて作業中に相談される者。主な相談内容は、検量モデルに関する技術的な課題である。割り当てる担当者の人数に制限は無く、0人または複数人としてもよい。以下の三つの役割についても、割り当てる担当者の人数に制限は無い。



4) **I**: 通知先。作業の完了が通知される者。主な通知手段は、報告書の回覧である。

**R** 及び **A** に割り当てる担当者の人数は各作業で唯一人に制限されるが、実際上は **R** 以外の担当者が作業を補助すること及び **A** 以外の担当者が確認者として文書に署名することがある。上記の四つの役割分類が不十分である場合、追加の役割分類を定義することができる。例えば、**RACI** の改良版として **RACI-VS**（確認者 **V** 及び署名者 **S** の追加）<sup>56)</sup> や **RASCI**（支援者 **S** の追加）<sup>57)</sup> が報告されている。これらを参考として、本検討では確認者（**V**）及び支援者（**S**）の二つの役割分類を追加した：

5) **V**: 確認者。作業を確認し、その作業に関する文書に署名する者。本役割分類を追加した理由は、メンテナンス関連文書は文書作成、確認、及び承認を経て最終化されるためである。

6) **S**: 支援者。**R** の指示下で補助的に作業を遂行する者。本役割分類を追加した理由は、実際上、複数の担当者が一つの作業を遂行することがあり得るためである。

本検討では **R**、**A**、**C**、**I**、**V**、及び **S** の六つの役割分類を使用する。IDEF0 に基づく作業に適用した **RACI** 表の概念図を Fig. 15 に示す。**RACI** 表は、Microsoft Excel Professional 2010（Microsoft）を用いて作図した。

	A1	A2	A3
<b>R</b> : Responsible (to perform)			
<b>S</b> : Supportive (to support R)			
<b>A</b> : Accountable (to sign as approver)	Subactivity 1	Subactivity 2	Subactivity 3
<b>V</b> : Verifier (to sign as verifier)			
<b>C</b> : Consulted (during the process)			
<b>I</b> : Informed (after the process)			
Assigned member 1	R	-	R
Assigned member 2	A	R	S
Assigned member 3	V	I	C

Fig. 15 Example of a RACI matrix

### 3.2.3 知識統合の効果を評価する指標

業務プロセスモデル化に基づいて検量モデルメンテナンスに関する知識を統合することで、検量モデルメンテナンス手順の理解を向上させ作業効率を向上させることが期待される。しかし、そうした手順の理解や作業効率を定量的に評価することは困難である。そこで、本検討ではそれらを実際の指標の一つとして、知識統合の前後における検量モデルメンテナンス **SOP** の内容及びメンテナンスの工数を用いることとした。

検量モデルメンテナンス **SOP** の内容の変化は、知識統合によって手順の理解が向上したことを反映している。これは、IDEF0 モデル及び **RACI** 表の作成を含む業務プ

ロセスモデル化で初めて得られた知識を SOP に組み込んだためである。得られた知識を SOP に組み込むことで、業務プロセスモデル化の専門家ではない担当者に対して得られた知識を共有することができる。まず、検量モデルメンテナンス SOP の内容の定性的な変化を列挙した。そして、検量モデルメンテナンス SOP の情報量をページ数に基づいて定量的に評価した。ページ数は、フォントサイズ及び字幅を揃えた上で、以下の規則に従い算出した。

- 1) 文章 37 行=1 ページ
- 2) 表 1 つ=1/2 ページ
- 3) 図 1 つ=1/2 ページ
- 4) メンテナンス業務のための計画書または報告書の様式 1 つ=1 ページ

メンテナンス工数は作業する担当者の人数に作業時間を乗じることで算出し、単位を person-hour とした。工数を正確に計算するために、各作業を報告書作成等の詳細な仕事に細分化し、各細分化された仕事について工数を算出した。例えば、報告書の作成に 1 人で 3 時間かかり、報告書の確認及び承認に 5 人で各 1 時間かかった場合、工数は  $1 \text{ person} \times 3 \text{ hours} + 5 \text{ people} \times 1 \text{ hour} = 8 \text{ person-hours}$  となる。詳細な仕事の工数を合計したものをメンテナンス工数と定義した。

### 3.3 結果及び考察

#### 3.3.1 検量モデルメンテナンスの枠組み

検量モデルメンテナンスとは、現行の検量モデルの推定精度を継続的に検証し、必要に応じてその頑健性を向上させるために検量モデルを更新することである。検量モデルを使用して推定値を算出する度に検量モデルメンテナンスを実施することは非現実的であるため、メンテナンスを実施する契機を設定することが必要である。従って、Fig. 16 に示すように、検量モデルメンテナンスの枠組みを「メンテナンスの必要性判断」及び「メンテナンス実施」の二つの部分に分けて作成した。本枠組みを作成するにあたり、各種の規制ガイドライン<sup>32), 41)</sup>及び USP<sup>33)</sup>を参考とした。

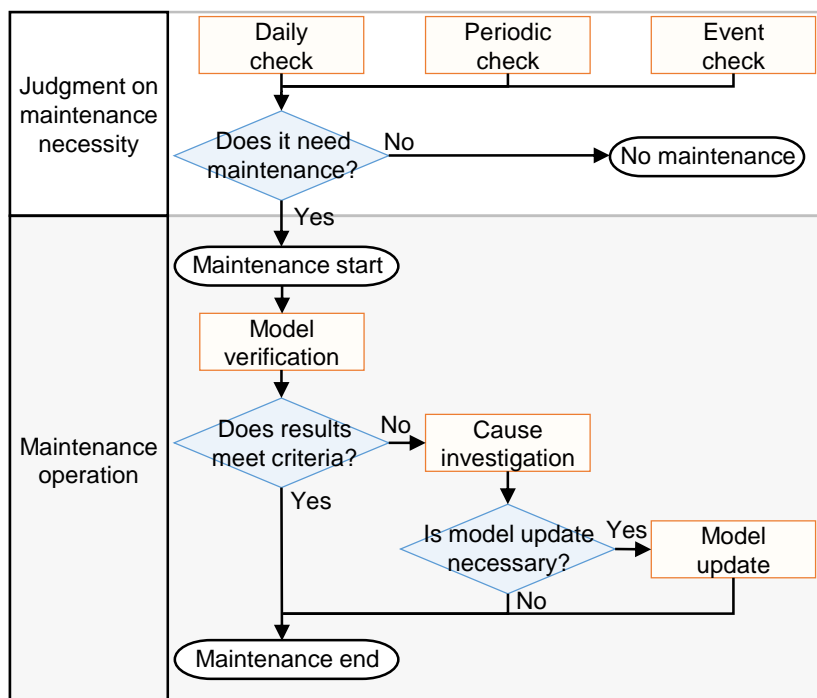


Fig. 16 The established framework of calibration model maintenance

### 3.3.1.1 メンテナンスの必要性判断

推定精度を検証するためには NIRS 及び参照法の両方で測定した値を比較することが必要となるが、日常的な生産では NIRS または参照法のいずれか一方を対象とする品質特性を評価する。通常は NIRS で対象を測定し、得られた NIR スペクトルを検量モデルに当てはめて推定値を算出する。NIRS を利用できない場合はその代替として参照法を用いる。従って、日常的には NIRS で得られた結果のみに基づいて検量モデルメンテナンスの必要性を判断する。本検討では、メンテナンスの必要性を判断するにあたり Table 12 に示す三つのチェック項目：日常チェック、定期チェック、及び事象チェックを採用した。

日常チェックに関しては、EMA<sup>32)</sup>及び ICH<sup>7)</sup>のガイドライン及び USP<sup>33)</sup>が外れ値の検出及びトレンド解析に基づく日常的なモニタリングを要求している。これらの要求に従い、NIRS を用いて測定を実施する際は NIR スペクトルの外れ値及び推定値の継時的傾向を評価することとした。

外れ値を評価する目的は、現行のキャリブレーションセットで保証する範囲外の NIR スペクトルを検出することである。例えば、何らかの原因でサンプル物性や測定装置の状態が変動した場合、得られる NIR スペクトルが著しく変動する可能性がある。この外れ値を検出するために、多変量統計工程解析 (multivariable statistical process control: MSPC) を用いることができる<sup>58), 59), 60)</sup>。MSPC は、検量モデルで説明可能な空間内における NIR スペクトルの変動及びその空間と NIR スペクトルとの距離の指標として、Hotelling's  $T^2$  及び Q 残差統計量を用いる。Hotelling's  $T^2$  または Q 残差統計

量が規定値を超えた場合は、その NIR スペクトルを外れ値と判断し、検量モデルメンテナンスが必要と判断する。

継時的傾向を評価する目的は、NIR 推定値が将来的に許容限界値を超える可能性を予測し、問題を未然に防ぐことである。例えば、測定装置の経年劣化や湿度の季節変化といった環境変化が NIR スペクトルに影響する可能性がある。NIR 推定値の継時的傾向を評価するために、Western Electric rule<sup>61)</sup>を用いることができる。Western Electric rule は、時系列の推定値に関する管理図及びその管理図を 3 つの領域に分ける管理幅を用いる。管理幅を設定する際は、製剤の品質管理戦略に基づいて許容される推定誤差を考慮する。各管理幅内の領域における NIR 推定値の数が規定値を超えた場合は、検量モデルメンテナンスが必要と判断する。

定期チェックに関しては、ICH ガイドライン<sup>7)</sup>が定期的なメンテナンスを要求している。この要求に従い、最後の検量モデルメンテナンスから半年間が経過した場合、検量モデルメンテナンスが必要と判断することとした。

事象チェックに関しては、EMA ガイドライン<sup>32)</sup>及び USP<sup>33)</sup>が Table 12 に示す項目の変更を管理することを要求している。これらの項目は製造条件及び測定条件に分類できる。これらの要求に従い、製造条件及び測定条件を変更した場合に、検量モデルメンテナンスが必要と判断することとした。変更し得る製造条件の例として、サンプリング器具、原料の製造元、原料の等級、製剤の処方、製造場所、製造装置、及び製造工程パラメータがある。これらの製造条件が変更されると、顆粒や錠剤等の測定サンプルの物理的性質が変動する可能性がある。それによって NIR スペクトル及び NIR 推定値が変動する可能性がある。変更し得る測定条件の例として、キャリブレーションセット、NIR 装置の光源、NIR 装置本体、測定場所、解析ソフトウェア、参照測定法、及び測定手順がある。これらの測定条件が変更されると、NIR スペクトル及び NIR 推定値が変動する可能性がある。

Table 12 Methods to judge maintenance necessity according to regulatory requirements

Check method	Detection target	Regulatory requirement of model maintenance timing	Reference
Daily check (outlier check)	Sudden variation in query data caused by unexpected problems such as human error in operation measurements or equipment failure	When outlier is detected	32), 33)
Daily check (trend check)	Gradual variation in query data caused by environmental change such as aging, deterioration of equipment, or seasonal variation of humidity	During routine manufacturing. Use monitoring such as trend analysis	7)
Periodic check	Deterioration in estimation accuracy, which is undetectable using only query data	Periodic maintenance	7)
Event check	Anticipated risk of deterioration of estimation accuracy by environmental changes such as equipment	When at least one of the following items has changed: - spectral library - equipment consumables with similar including lamps, sampling devices, location and software upgrades - physical properties of the material - source of material supply - critical attribute(s) of material(s) - composition of the test sample or finished product - manufacturing process - sources or grades of raw materials - reference analytical method - instrument hardware	32), 33)

### 3.3.1.2 メンテナンス実施

メンテナンス実施の作業は、モデル検証、原因追究、及びモデル更新に大別される。まずモデル検証において検量モデルの推定精度を検証する。検証結果が所定の基準を満たした場合はそのメンテナンスを終了する。一方、検証結果が所定の基準を外れた場合はその原因を追究する。原因追究の結果から検量モデルの更新が不要と判断されれば、そのメンテナンスを終了する。一方、原因追究の結果から検量モデルの更新が必要と判断されれば、その推定精度を向上させるために検量モデルを更新する。

モデル検証について、EMA ガイドライン<sup>32)</sup>は以下のように述べている:「NIRS 及び参照法を用いて同じサンプル群を測定した結果を比較することが (可能であれば) NIRS の検証の一部となる」。また、検量モデルの最も重要な性能指標は推定精度である。従って、モデル検証では推定精度を評価対象とした。推定精度の評価では、NIRS 及び参照法で試験サンプルを測定し、得られた推定値及び参照値を直接比較する。推

定精度を SEP で評価する<sup>32)</sup>。SEP が所定の基準値以内であった場合、その推定精度は問題ないと判定される。一方、SEP がその基準値を超えた場合、その推定精度は低下したと判定され、その原因が追究される。SEP の基準値を設定する際は、製剤の品質管理戦略に基づいて許容可能な推定誤差を考慮する。例えば、実際の推定誤差が基準値（許容可能な推定誤差の最大値）に等しい場合でも最終的な製品品質を保証できることが最低限必要な水準と考えられる。

原因追究について、USP は以下のように述べている:「性能が許容できない場合は是正措置が必要である。その際、性能が変化した原因を特定するための調査を実施する。その結果、NIRS を継続して使用することが不適切とされる場合もある」。この記載に従い、モデル検証の結果が不適となった場合は、推定精度が低下した原因を追究する。

原因追究では、測定装置、測定操作、及び物質特性を逐次的に調査する。原因追究の流れを Fig. 17 に示す。

測定装置の調査では、その製造元に点検を依頼する。推定精度が低下した原因を測定装置の故障に帰結できる場合は、測定装置を修理し、改めてモデル検証を実施する。モデル検証の結果が適合となる場合は、モデル更新が不要と判断する。モデル検証の結果が不適となる場合は、測定装置の状態変化に対する頑健性を向上させるためにモデル更新が必要と判断する。推定精度が低下した原因を測定装置の故障に帰結できない場合は、測定操作を調査する。

測定操作の調査では、測定者への聞き取り調査を実施するとともに、測定 SOP を精査する。推定精度が低下した原因を測定操作に帰結できる場合は、同様の事象を繰り返さないように測定 SOP を改訂し、改めてモデル検証を実施する。モデル検証の結果が適合となる場合は、モデル更新が不要と判断する。モデル検証の結果が不適となる場合並びに推定精度が低下した原因を測定装置及び測定操作に帰結できない場合は、測定サンプルの物質特性を調査する。

物質特性の調査では、リスク評価を実施して物質特性のリスク因子を特定し、そのリスク因子が推定精度に与える影響を検証する。推定精度が低下した原因を物質特性に帰結できる場合は、その物質特性の変動が製剤品質に与える影響を検証する。その物質特性の変動が製剤品質に影響を与えないことが確認できた場合には、その変動に対する頑健性を向上させるためにモデル更新が必要と判断する。

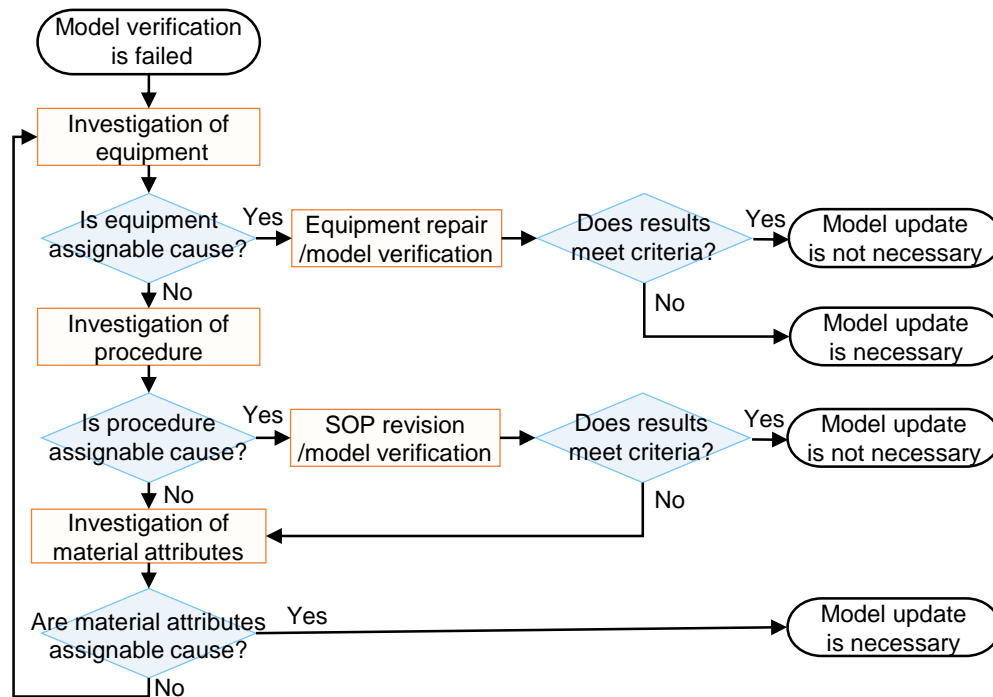


Fig. 17 The workflow of cause investigation

モデル更新では、特定した原因に由来する変動に対する頑健性を向上させるために、検量モデルを更新する。検量モデルを更新する際は、現行の検量モデルで採用しているキャリブレーションセット、スペクトル前処理、及び波数領域を変更することで、更新後の検量モデルとする。モデル更新時にも第一章及び第二章で提案した波数選択法を用いることができる。本波数選択法を用いることで、キャリブレーションセット及びスペクトル前処理に応じた波数選択結果が得られる。事例として、推定精度が低下した原因をサンプルの水分含量の変動に帰結できた場合を考える。この場合、水分含量の変動に対する頑健性を向上させるために以下の対策が考えられる: 1) キャリブレーションセットに水分含量の異なるサンプルを追加する、2) 現行の検量モデルで採用している波数領域のうち、水分含量の変動がスペクトル強度に影響を与える波数領域を除外する、及び 3) 水分含量の変動が NIR スペクトルに与える影響を最小化するようなスペクトル前処理を選択する。更新後の検量モデルを構築した後、分析法バリデーションを実施して更新後の検量モデルの性能を検証する。Nakagawa ら<sup>62)</sup>は詳細なバリデーション項目を含む実践的なバリデーション手順を示している。

### 3.3.2 業務プロセスモデル化に基づく知識統合

#### 3.3.2.1 知識統合前の検量モデルメンテナンス

Fig. 16 に示す検量モデルメンテナンスの枠組みに基づき、最初のメンテナンス SOP を作成した。商用生産工場にて、本 SOP に基づく検量モデルメンテナンスが三年間運用された。当該工場において、検量モデルメンテナンスは新規の運用であった

ため、経験的な知識は限られていた。また、検量モデルメンテナンスは GMP に従って複雑な文書管理を伴うため、最初の SOP に様々な状況に対応する詳細な手順を定義することは困難であった。そこで、その状況に応じて詳細な手順を確認するために、その都度品質保証責任者等の複数の担当者との打合せを実施した。このように状況毎に繰り返し打合せを実施すれば運用の信頼性を保証できるが、必要以上に作業負担を増大させる可能性があった。また、最初の SOP に記載された役割と責任の内容が簡素であったため、情報を伝えるべき相手を確認するために別の担当者に相談する等、情報伝達が遅延する可能性があった。以上の状況から、運用の信頼性を維持したまま作業負担を改善する必要があった。

### 3.3.2.2 業務プロセスモデルの作成

検量モデルメンテナンスの作業効率を改善するために、検量モデルメンテナンスにおける情報の流れ、並びに役割及び責任を業務プロセスモデルである IDEF0 モデル及び RACI 表に整理した。この業務プロセスモデル化の範疇は、Fig. 16 に定義した検量メンテナンスの枠組みにおけるメンテナンス開始から終了までとした。

検量モデルメンテナンスにおける主要な課題の一つに技術移転がある<sup>30)</sup>。本検討の場合、研究所が検量モデルメンテナンスの枠組みを設計し、その枠組みを商用生産工場に技術移転した。技術移転の後、工場は検量モデルメンテナンスを運用化した。このメンテナンスの運用を整理する上で、研究所と工場の双方が有する知識を統合することが有用である。その知識を統合するために、研究所と工場とが協業して業務プロセスモデル化を実施した。研究所は、メンテナンス担当者及び品質保証責任者等の工場の担当者との打合せを数回実施し、その打合せを踏まえて IDEF0 モデル及び RACI 表を作成した。各打合せにおいて、IDEF0 モデル及び RACI 表を作成しながら検量モデルメンテナンスにおける情報の流れを議論した。必要に応じて、各打合せ終了後に工場の担当者が各部署の責任者から必要な情報を収集した。

作成した IDEF0 モデル及び RACI 表を Fig. 18 及び Fig. 19 に示す。IDEF0 モデルでは、最上位の作業を「メンテナンスを実施する (A0)」と定義した。検量モデルメンテナンスが必要と判断された場合、メンテナンス指示書が発行され、これがメンテナンスを開始する契機となる。本メンテナンスでは、必要に応じて現行の検量モデルを更新し、更新後の検量モデルを構築する。任命された担当者が、SOP、最初のキャリブレーションセット、新規サンプル、新規サンプルに関する情報、及び装置を用いてメンテナンスを実施する。また、担当者を RACI 表で定義した。「メンテナンスを実施する (A0)」作業を 5 つの小作業: 「メンテナンスを管理する (A1)」、「モデルを検証する (A2)」、「モデルを更新し、妥当性を確認する (A3)」、「メンテナンス履歴を更新する (A4)」、及び「資源を供給する (A5)」に細分化した。RACI を用いてこれらの A1 から A5 の小作業に関する役割及び責任を定義した。



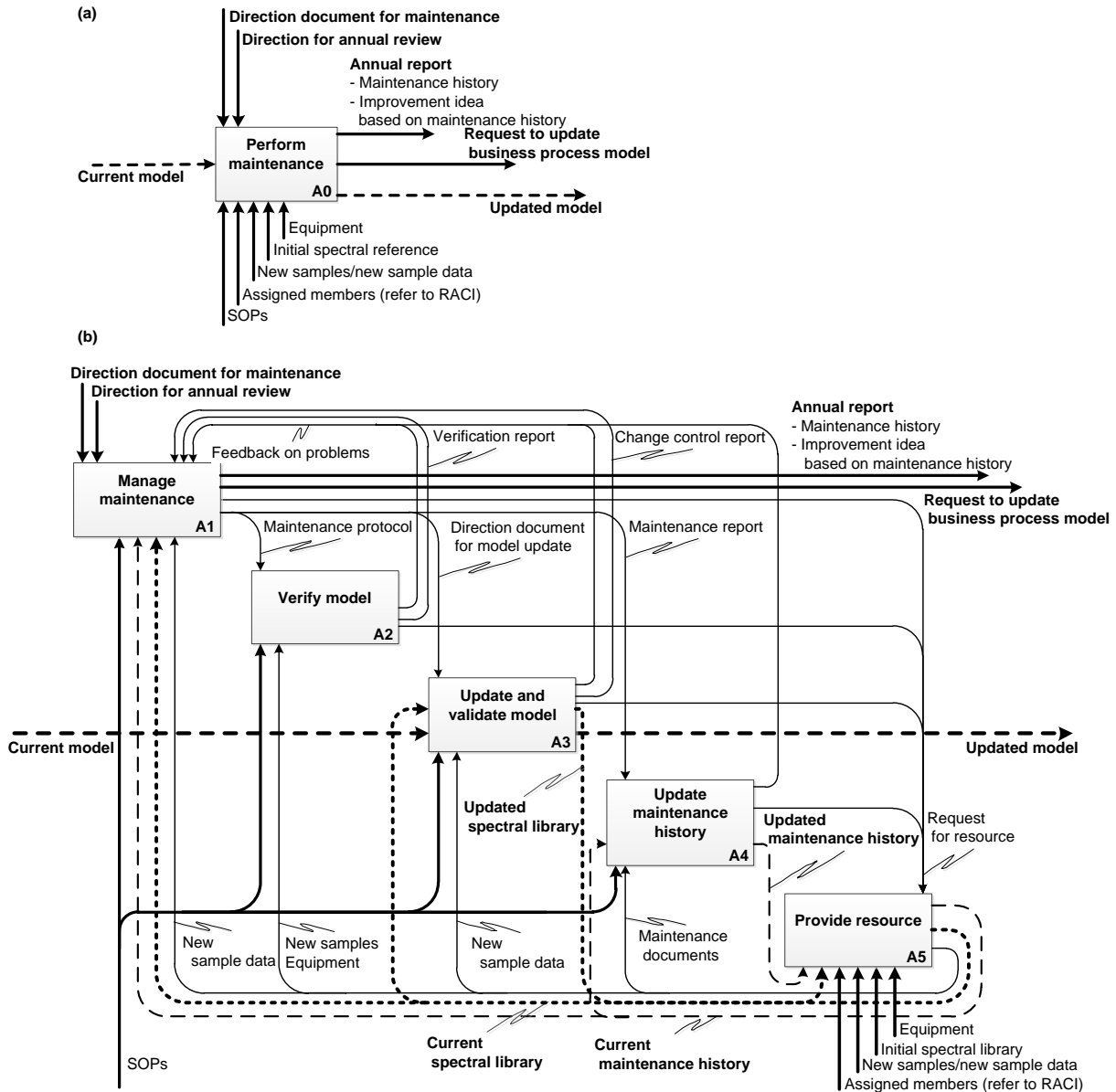


Fig. 18 The IDEF0 model of calibration model maintenance. (a) Top activity (A0), (b) Subactivities (A1 to A5). The bold arrows represent elements that are commonly used in the top activity and the subactivities. Each kind of dashed line represents the corresponding items being updated in model maintenance

	A1	A2	A3	A4	A5
R: Responsible (to perform)	Manage maintenance	Verify model	Update and validate model	Update maintenance history	Provide resource
S: Supportive (to support R)					
A: Accountable (to sign as approver)					
V: Verifier (to sign as verifier)					
C: Consulted (during the process)					
I: Informed (after the process)					
Maintenance staff	R	I	R	R/A <sup>a)</sup>	R/A <sup>a)</sup>
Manufacturing control manager	V	I	V	-	-
Quality control manager	V	A	I	-	S
Quality control operator	-	R	I	-	-
Quality assurance manager	A	-	A	-	-
Manufacturing technology staff	C	-	C	C	-

Fig. 19 The RACI matrix of model maintenance. Roles were assigned to the subactivities (A1 to A5) of the IDEF0 model. a: The member doubles as R and A and approves the process without a signature because the process produces no document

### 3.3.2.3 IDEF0 モデルの解釈

「メンテナンスを管理する (A1)」では、検量モデルメンテナンス SOP に従い以下の 5 つの小作業を実施する:

#### 1) メンテナンスを開始する

日常チェック、定期チェック、または事象チェックに基づいてメンテナンスが必要と判断された場合、A1 はメンテナンス指示書を受け取る (A1 への制御の矢印)。メンテナンス業務を開始するために、A1 はメンテナンス計画書を作成し、A2 へ提出する (A1 から A2 への制御の矢印)。

#### 2) 原因を追究し、モデル更新の必要性を判断する

モデル検証の結果が所定の基準を外れた場合、A1 は A2 からモデル検証報告書を受け取る (A2 から A1 への制御の矢印)。そして、A1 は新規サンプルに関する情報、現行のキャリブレーションセット、及び測定 SOP (A1 への機能の矢印) を用いてモデル検証の結果が基準を外れた原因を追究する。原因追究の結果に基づいて A1 はモデル更新の必要性を判断し、原因追究の報告書を作成する。モデル更新が必要と判断した場合、A1 はモデル更新指示書を作成し、A3 へ提出する (A1 から A3 への制御の矢印)。モデル更新が不要と判断した場合、メンテナンス業務を終了する。

#### 3) メンテナンスを終了する

モデル検証の結果が所定の基準を満たした場合、A1 は A2 からモデル検証報告書を受領する (A2 から A1 への制御の矢印)。モデル検証の結果が所定の基準を外れた

が、モデル更新が不要と判断された場合、A1 は 2)において原因追究の報告書を作成済である。例えば測定装置の異常や測定操作の誤りが原因であった場合がこれに相当する。モデル検証の結果が所定の基準を外れ、かつ検量モデルが更新された場合、A1 は A3 から変更管理報告書を受領する (A3 から A1 への制御の矢印)。いずれの場合においても、A1 はモデル検証の報告書、原因追究の報告書、及び変更管理の報告書を取りまとめたメンテナンス報告書を作成し、A4 へ提出する (A1 から A4 への制御の矢印)。

#### 4) 年次照査を実施する

検量モデルメンテナンスの結果は毎年照査される。A1 が年次照査の指示 (A1 への制御の矢印) を受けた場合、A1 は現行のメンテナンス履歴を用いて検量モデルメンテナンスに関する年次照査を実施する。そして、A1 は年次照査の報告書を作成し、品質保証責任者に提出する。

#### 5) 問題に対応する

他の小作業において装置の故障や担当者の不在といった問題が発生した場合、A1 は当該の小作業から問題の報告を受ける (当該の小作業から A1 への制御の矢印)。その報告を受けて、A1 は適時その問題に対応する。必要に応じて、A1 は IDEF0 及び RACI に基づく業務プロセスモデルを更新することを研究所に依頼する。

「モデルを検証する (A2)」では、検量モデルメンテナンス SOP に従い以下の 2 つの小作業を実施する:

##### 1) 新規サンプルを測定し、推定精度を評価する

A2 が A1 からメンテナンス計画書を受け取った場合 (A1 から A2 への制御の矢印)、A2 は NIRS 及び参照法の両方で新規サンプルを測定する。

##### 2) 推定精度を評価する

新規サンプルを測定した後、A2 は NIR 推定値と参照値との SEP を算出し、その結果が所定の基準に適合するか否かを判定する。その結果に基づき、A2 はモデル検証の報告書を作成し、A1 へ提出する (A2 から A1 への制御の矢印)。

「モデルを更新し、その妥当性を確認する (A3)」では、検量モデルメンテナンス SOP に従い以下の 4 つの小作業を実施する:

##### 1) モデル更新を計画する

モデル更新が必要と判断され、A3 が A1 からモデル更新指示書を受け取った場合 (A1 から A3 への制御の矢印)、A3 はモデル更新計画書を作成する。

##### 2) 更新後の検量モデルを構築する

モデル更新計画書に従い、A3 は現行のキャリブレーションセット及び新規サンプルに関する情報 (A3 への機能の矢印) を用いて更新後の検量モデルを構築する。現行の検量モデルに加えて、現行のキャリブレーションセットも更新される。従って、

現行の検量モデル及びキャリブレーションセットは入力（矢印）で、更新後の検量モデル及びキャリブレーションセットは出力（矢印）で示される。

### 3) 変更管理を計画する

現行の検量モデルから更新後の検量モデルへの変更は、GMP 管理上の変更管理として実施される。従って、A3 は変更管理計画書を作成する。

### 4) 変更管理を実施する

変更管理計画書に従い、A3 は更新後の検量モデルを製造管理システムに登録する。そして、A3 は変更管理報告書を作成し、A1 へ提出する（A3 から A1 への制御の矢印）。

「メンテナンス履歴を更新する（A4）」を、IDEF0 に基づく議論の過程で新規に作成した。これは知識管理の観点から、継続的なメンテナンス業務で得られる経験知識を製造工程の継続的改善に活用すべきと判断したためである。知識管理とは、製造工程、処方、及び製品に関する情報を収集し、解析し、保管し、広めるための体系的な取り組みである<sup>5)</sup>。知識管理の目的は、継続的改善を推進し、管理された状態を維持することである。検量モデルをメンテナンスするのは、推定精度が低下した危険性があると判断した場合である。この危険性は、NIR 推定値の外れ値や傾向異常に基づいて評価される。このような外れ値や傾向異常は、製造プロセス特性の変動を反映している可能性がある。従って、検量モデルメンテナンスが実施された履歴を評価することで、製造プロセス特性の変動をより早い段階で検出できる可能性がある。例えば、NIR 推定値の傾向異常が原因でメンテナンスが頻繁に発生していた場合、その製造プロセス特性が傾向を持って変動している可能性があると予想できる。このように検量モデルメンテナンスにより、単に検量モデルの推定精度を検証するだけでなく、製造プロセス特性に関する間接的な情報を得ることができる。こうした考えから、検量モデルメンテナンス業務で得られた知識を蓄積するために、メンテナンス履歴を新規に作成した。本履歴は検量モデルメンテナンスを実施する度に更新され、年次照査の折に参照される。本履歴を継続的に更新し、定期的に照査することで、検量モデルメンテナンスを製造プロセスに関する知識管理及び継続的改善に活用できると期待される。IDEF0 において、経験知識を適切に管理するために「メンテナンス履歴を更新する（A4）」を定義した。ただし、本履歴は Fig. 16 に示すメンテナンスの枠組みには定義されていない。A4 が A1 からメンテナンス報告書を受領した場合（A1 から A4 への制御の矢印）、A4 はメンテナンス関連文書（A4 への機能の矢印）に基づいて現行のメンテナンス履歴に新規のメンテナンス記録を追記する。結果として、A4 は入力（矢印）で示される現行のメンテナンス履歴を更新し、出力（矢印）で示される更新後のメンテナンス履歴を作成する。改訂後の検量モデルメンテナンス SOP に新たな項目としてメンテナンス履歴に関する手順を追記したため、A4 は改訂後の SOP に従って実施される。

「資源を供給する（A5）」では、有限の資源または原本となる資源を取り扱う。

例えば、適時使用のために予約を必要とする測定装置及び版管理を必要とする現行のキャリブレーションセットがこれにあたる。A5 が他の小作業の担当者から資源供給の依頼を受けた場合（A5 への制御の矢印）、A5 はその小作業の担当者に資源の使用を許可する。一方、無制限に使用できる資源に関しては、A5 に対する資源供給の依頼は不要である。例えば、検量モデルメンテナンス SOP は電子的に複写できるため、各小作業の担当者は A5 を介さずに SOP を読むことが出来る。

### 3.3.2.4 RACI 表の解釈

IDEF0 は各小作業において誰が何の役割を担うかを定義しない。そこで、RACI を用いて各小作業の役割及び責任を定義した。以下に示す任命された担当者が検量モデルメンテナンス業務に関与する：

- 1) メンテナンス担当者。検量モデルメンテナンス業務を遂行する。本担当者は、一連の作業の進捗を管理し、メンテナンス関連文書を作成する。
- 2) 製造管理責任者。製剤の製造を管理する。製造工程における PAT の手段として検量モデルを利用するため、本責任者は検量モデルメンテナンスについて確認者として関与する。
- 3) 品質管理責任者。モデル検証を管理する。本責任者は品質管理担当者に NIRS 及び参照法による測定を含むモデル検証試験を指示し、その試験結果を承認する。
- 4) 品質管理担当者。品質管理責任者の指示を受けて、NIRS 及び参照法による測定を含むモデル検証試験を実施する。
- 5) 品質保証責任者。製品品質の保証及び GMP 遵守について責任を負う。製品品質を推定するために検量モデルを利用するため、本責任者は検量モデル及びそのメンテナンス工程の妥当性を承認する。
- 6) 製造技術担当者。必要に応じて技術的な問題に対応する。本担当者は、検量モデルメンテナンス、NIRS、及び多変量解析を含む検量モデルに関する専門的な知識を有する。

Fig. 19 に A1 から A5 の小作業に関する役割及び責任を定義した RACI 表を示す。各小作業における代表的な役割を以下に述べる。

「メンテナンスを管理する (A1)」において、メンテナンス担当者 (R) はメンテナンス報告書を作成する実務を担う。製造管理責任者及び品質管理責任者 (V) は報告書を確認し、押印する。品質保証責任者 (A) は確認済の報告書を承認し、押印する。必要に応じて、製造技術担当者 (C) は検量モデルに関する技術的な問題に関する相談を受ける。

「モデル検証 (A2)」において、品質管理担当者 (R) は NIRS 及び参照法で新規サンプルを測定する実務を担う。品質管理責任者 (A) は測定結果を承認し、押印する。承認された測定結果がメンテナンス担当者及び製造管理責任者 (I) に展開され、

本小作業が完了したことが通知される。

「モデルを更新し、妥当性を確認する (A3)」において、メンテナンス担当者 (R) は変更管理報告書を作成する実務を担う。製造管理責任者 (V) は報告書を確認し、押印する。品質保証責任者 (A) は確認済の報告書を承認し、押印する。必要に応じて、製造技術担当者 (C) はモデル更新技術に関する相談を受ける。承認された報告書が品質管理責任者及び品質管理担当者 (I) に展開され、本小作業が完了したことが通知される。

「メンテナンス履歴を更新する (A4)」において、メンテナンス担当者 (R/A) はメンテナンス履歴を更新する実務を担う。本小作業では文書を作成しないため、押印すること無しにメンテナンス担当者 (R/A) 自身が更新したメンテナンス履歴を承認する。必要に応じて、製造技術担当者 (C) はメンテナンス履歴の記載内容について相談を受ける。メンテナンス履歴は年次照査等で必要となった際に参照されるものであるため、本小作業の終了を逐次通知する必要はない。従って、通知先 (I) は不要とした。

「資源を供給する (A5)」において、メンテナンス担当者 (R/A) は、他の小作業からの依頼を受けて現行のキャリブレーションセット等の資源を供給する。必要に応じて、品質管理責任者 (S) に測定装置等の測定に関する資源を供給することを指示する。本小作業では文書は作成しないため、押印すること無しにメンテナンス担当者 (R/A) 自身が資源の使用を承認する。A5 は簡潔な業務であるため、相談先 (C) 及び通知先 (I) は不要とした。

### 3.3.3 検量モデルメンテナンスの運用化

業務プロセスモデルを作成する過程で、各担当者はメンテナンス業務に関する理解を深めるとともに、詳細な手順に関する各自の知識を IDEF0 モデル及び RACI 表に統合した。この IDEF0 モデル及び RACI 表を参照することで、詳細な手順を確認することができる。しかし、業務プロセスモデル化の専門家ではない担当者にとって IDEF0 モデルを理解することは困難である。そこで、IDEF0 モデルに統合された知識を理解しやすくするために、IDEF0 モデルの情報を文章に書き換えた。また、どの担当者でも参照できるように、検量モデルメンテナンス SOP にその文章を追記し、その SOP を更新した。本事例では、IDEF0 モデルから文章への書き換えは研究所が担当した。商用生産工場において、更新された SOP に従って検量モデルメンテナンスを運用することとした。

#### 3.3.3.1 検量モデルメンテナンス SOP の更新

各担当者は検量モデルメンテナンス SOP に基づいてメンテナンスに関する詳細な手順を理解し、それに従ってメンテナンス業務を遂行する。従って、メンテナンス業務を効率的に実施する上で SOP に適切な情報が記載されていることが重要である。

SOP 以外に担当者の実務経験、教育、及び個人の能力等もメンテナンス業務の効率に影響を与えるが、SOP を適切に更新することで各担当者のメンテナンス業務に対する理解を向上させることができる。本検討において業務プロセスモデル化に基づいて知識を統合した結果、メンテナンス業務に関する詳細な手順並びに必要な役割及び責任が明確化された。また、統合した知識を踏まえて SOP を改訂した。SOP の内容及び情報量の変化を Table 13 及び Fig. 20 に示す。知識統合を踏まえた SOP の改訂により、SOP の情報量は約 2 倍に増加した。

最も情報量が増加したのは各作業を開始する契機及び各作業で作成される成果物等のメンテナンス手順に関する内容であった。これらの手順に関する記載を SOP に追記したため、どの担当者でも SOP を参照することで詳細な手順を理解することができる。さらに、改訂後の SOP に Fig. 16 に示すメンテナンスの枠組みを追記した。メンテナンスの枠組みに相当する情報量（ページ数）は少ないが、担当者全員がメンテナンス手順の全体像を理解する上で重要な内容である。手順に関する理解が向上したことで、GMP 管理下におけるメンテナンス業務をより迅速に遂行できるようになった。

手順に加えて、追記された役割、より詳細な責任に関する記載、及び RACI 表等の役割に関する情報量が増加した。例えば、品質管理責任者は品質管理担当者に検証試験を指示し、その結果を承認するという間接的な役割であったため、最初の SOP にその役割が記載されていなかった。しかし、RACI 表に基づいて役割分担に関する知識を統合した結果、このような間接的な役割が明確化されたため、改訂後の SOP に品質管理責任者に関する記載が追記された。さらに、役割及び責任の概要として RACI 表が追記された。こうした役割に関する記載が増加したことで、担当者自身の役割だけでなく他の担当者の役割に関する理解を深めることができた。このような役割に関する理解が向上したことで、より適時適確な情報伝達が可能となった。

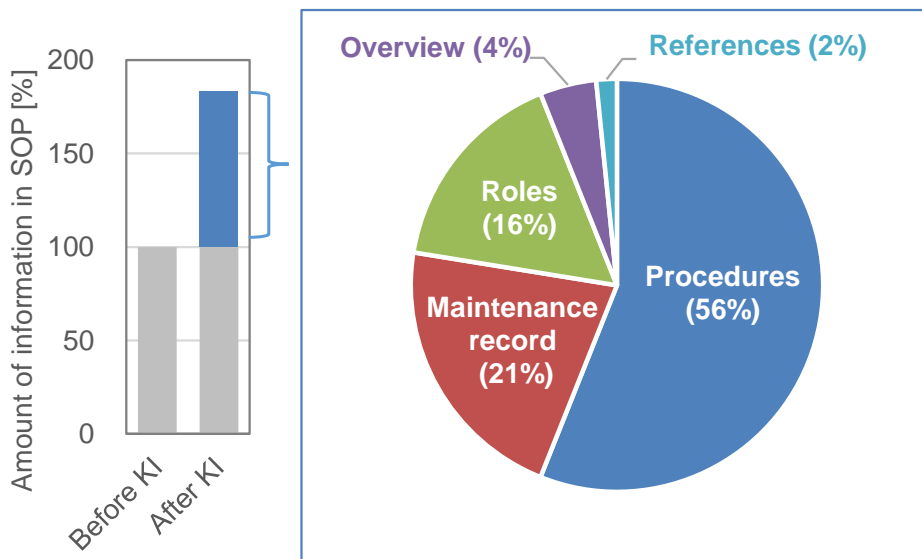


Fig. 20 Increased amount of information in model maintenance SOP after knowledge integration (KI). Left: the amount of information before and after KI. Right: classification of the increased amount of information according to its content

Table 13 List of changes in content of model maintenance SOP through knowledge integration

Added content	Example
<b>Procedure</b>	
Trigger and output items	At the start of model maintenance, a direction document for maintenance is issued. Prepare a maintenance protocol.
Documentation	When the model is updated, prepare the change control protocol and report in accordance with a change control SOP
<b>Maintenance record</b>	
Section of maintenance record	When model maintenance is finished, update the model maintenance record.
Maintenance record form	Maintenance periods and cause of maintenance
<b>Roles</b>	
Additional roles	Quality control manager
Detailed responsibilities	The person who is required to direct the quality control operator conducting a verification test, and who approves the verification test results
A RACI matrix	Fig. 19
<b>Overview</b>	
Model maintenance framework	Fig. 16
<b>References</b>	
Related SOPs	A change control SOP

### 3.3.3.2 作業効率の改善

實際上、検量モデルメンテナンスを運用する際はその工数が重要な課題となる。本検討では、知識統合及び検量モデルメンテナンス SOP 更新を実施した前後におけるメンテナンス業務に要する工数を定量的に比較した。



IDEF0 に基づいて知識を統合したことで、想定されるあらゆる状況に対する手順を明確化することができた。また、業務プロセスモデル化の過程で、あらゆる状況に対する手順を模擬的に実施し、その手順通りに作業を遂行できることを検証した。この検証された手順を改訂後の SOP に追記した。本 SOP を参照することで適切な手順を確認できる。これによって、SOP 改訂後は適切な手順を確認するための打合せを実施する回数を低減することができた。以上のように、改訂後の SOP を用いることで、検量モデルメンテナンスをより迅速に適切な手順で遂行することが可能となった。また、RACI に基づいて知識を統合したことで、GMP 管理下でのメンテナンス業務に関する必要十分な役割分担を明確化することができた。その結果、知識統合前の役割分担には必要以上の確認者が含まれていることが判明したため、その確認者を知識統合後の役割分担から除外した。確認者の人数を削減したことで、文書作成から承認までの確認に必要な工数及び時間を低減することができた。

検量モデルメンテナンスの工数はモデル更新の有無によって大きく変動する。そこで、検量モデルメンテナンスを 2 つの部分：(a)「モデル検証に相当する部分。但しメンテナンス開始、終了、及び履歴更新を含む」並びに (b)「モデル更新に相当する部分。但し (a) の部分を除外し、原因追究を含む」に分割して検討した。Fig. 16 に示す検量モデルメンテナンスの枠組みを参照すると、モデル検証の結果が適合した場合は (a) の部分のみが実施される。一方、モデル検証が不適となりモデル更新が実施される場合は、(a) 及び (b) の部分が実施される。各部分の工数を算出し、知識統合前後の工数を比較した結果を Fig. 21 に示す。知識統合を実施したことによって、(a) の部分は 14 person-hours の工数が削減され、(b) の部分は 59 person-hours の工数が削減された。

事例研究として、(a) の部分における手順及び工数の変化を以下に述べる。知識統合前は、メンテナンス担当者はメンテナンス計画書及び報告書を作成するにあたり品質保証責任者等の他部署の責任者との事前打合せを実施していた(事前確認の工数: 4 person-hours)。また、白紙の状態から数ページのメンテナンス計画書及び報告書を作成していた(文書作成の工数: 14 person-hours)。さらに、メンテナンス計画書及び報告書は 5 人の担当者によって確認された後、1 人の担当者によって承認される手順となっていた(文書押印の工数: 14 person-hours)。知識統合の過程で、想定されたあらゆる状況に対する手順が明確化された。その結果、メンテナンス計画書及び報告書の様式を作成し、想定される状況を予め記載することができた。本様式は 1 ページの分量であり、より簡便にメンテナンスの結果を記載することができる。知識統合後は、メンテナンス担当者はメンテナンス計画書及び報告書をより迅速に作成することができ(文書作成の工数: 7 person-hours。7 person-hours の削減)、他部署の担当者との事前打合せに要する時間が低減された(事前確認の工数: 2 person-hour。2 person-hours の削減)。さらにメンテナンス計画書及び報告書は 2 人の担当者によって確認された後、1 人の担当者によって承認される手順となった(文書押印の工数: 8 person-hours。

6 person-hours の削減)。これはメンテナンス計画書及び報告書用の様式に予め記載されている内容が、知識統合の打合せ時に複数の責任者によって確認済であるためである。メンテナンス履歴を更新することが新たな手順として追加されたが、それによって増加した工数はメンテナンス記録の概要をデータベースに入力するのみの僅かな工数である（メンテナンス履歴更新の工数: 1 person-hour）。以上をまとめると、知識統合に基づくメンテナンス手順の改善によって、GMP 管理下における運用の適切性を維持した上でメンテナンスの (a) の部分を一回実施するにあたり  $7+2+6-1=14$  person-hours の工数を削減することができた。

(b) の部分についても同様に、知識統合に基づく手順の改善によって、事前確認について 25 person-hours の工数及び文書押印について 16 person-hours の工数を削減することができた。さらに、知識統合によって、モデルを更新しその妥当性を確認する手順が明確化された。その結果、更新後の検量モデルの性能を確認するための様式を作成し、評価用の計算式を予め組み込むことができた。本様式に測定結果を入力することで、解析結果を自動的に算出することができる。従って、知識統合後は、メンテナンス担当者は検量モデルをより迅速に更新し、その妥当性を確認することができるようになった（18 person-hours の削減）。以上をまとめると、知識統合に基づく手順の改善によって、GMP 管理下における運用の適切性を維持した上でメンテナンスの (b) の部分を一回実施するにあたり  $25+16+18=59$  person-hours の工数を削減することができた。なお、上記の評価ではモデル更新時の波数選択を試行錯誤で実施すると仮定して工数を算出した。第一章及び第二章で提案した波数選択法（自動的にかつ従来法よりも短い計算時間で波数選択を完了できる）を用いるため、もしこれらの波数選択法を利用すればさらなる工数削減が期待できる。なお、これらの波数選択法に要する計算時間はそのモデル更新に使用するキャリブレーションセットのデータ量（サンプル数及び NIR スペクトルの波数点の数）及び計算機の性能に依存するため、本評価の対象外とした。

検量モデルメンテナンスは製品ライフサイクルを通じて継続的に実施されるため、長期的な工数を評価することとした。本事例における製剤では、その品質管理戦略の中で 4 つの検量モデルを採用しており、検量モデル毎にメンテナンスを実施する。各検量モデルについて、少なくとも半年に 1 回は定期チェックに基づくメンテナンスを実施する。つまりメンテナンスを実施する頻度は少なくとも 4 つの検量モデル  $\times$  2 回/年 = 8 回/年となる。各メンテナンスにおいて、モデル検証は必ず実施されるが、モデル更新は必要に応じて実施される。モデル更新を実施する頻度を予測することは困難であるので、モデル更新を年間 0 回、1 回、または 2 回実施する場合を想定することとした。知識統合に基づく手順の改善によって得られる工数の削減効果は、モデル更新が年間 0 回の場合は 112 person-hours/year、モデル更新が年間 1 回の場合は 171 person-hours/year、そしてモデル更新が年間 2 回の場合は 230 person-hours/year と計算された。しかしその一方で、知識統合に際して業務プロセスモデルを作成し、検量モ

デルメンテナンス SOP を更新するための工数を必要とした。この工数を回顧的に算出した結果、1年間の期間をかけて 209 person-hours の工数を消費していた。その工数の内訳は、業務プロセスモデルの作図（1人の担当で実施、30 person-hours）、業務プロセスモデルを用いたメンテナンス手順の議論（6~10人の担当で実施、計 154 person-hours）、及び SOP の更新（6~9人の担当で実施、計 25 person-hours）である。

削減された工数の累積として、知識統合を実施するために必要とされた工数及び知識統合によって削減された工数をまとめた結果を Fig. 22 に示す。削減された工数を最も小さく見積もった場合、つまりモデル更新を実施する頻度が年間 0 回であったと仮定しても、知識統合後 2 年間の運用で削減された工数は知識統合を実施するために必要とした工数を上回ることになる。

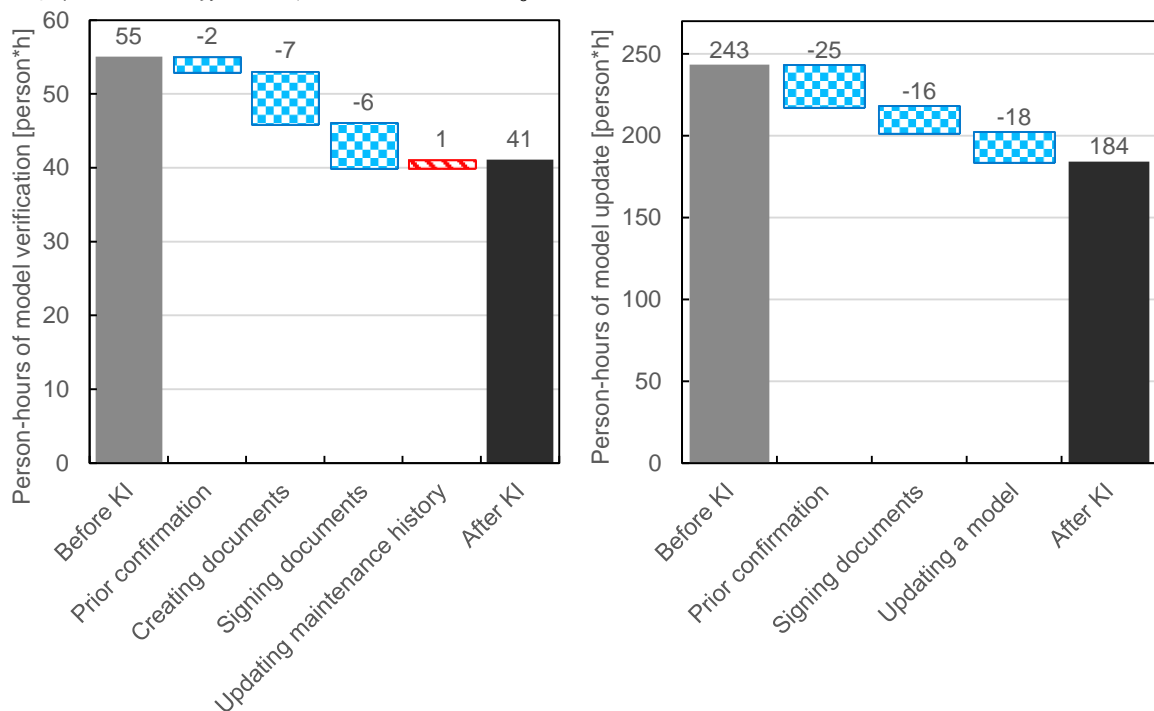


Fig. 21 Changes in person-hours of model maintenance by knowledge integration (KI) and the breakdown. The checkered columns represent decreased person-hours. The striped column represents increased person-hours. Left: Person-hours of model verification. Right: Person-hours of model update

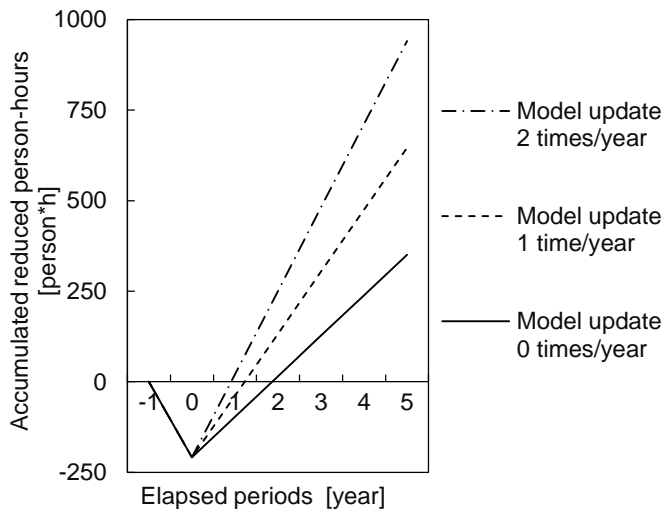


Fig. 22 Estimation of accumulated reduced person-hours of model maintenance. In the period from -1 year to 0 years, person-hours were dedicated to develop the business process model and to update the maintenance SOP. At the time point of 0 years, the updated SOP was enacted. After the time point of 0 years, person-hours will be reduced by the updated SOP. The estimation assumed that the model update is to be conducted 0, 1, or 2 times/year

### 3.3.3.3 検量モデルメンテナンス運用の検証

本事例における商用生産工場では、対象とする製剤について3つの定量分析用の検量モデル及び1つの定性分析用の検量モデルの計4つの検量モデルを採用している。知識統合及び検量モデルメンテナンス SOP の改訂後に、製造工程のプロセスバリデーションを実施し、各検量モデルを対象としてメンテナンス業務の手順を検証した。

本プロセスバリデーションでは、製剤3ロットを製造し、各ロットの製造において計画的に検量モデルメンテナンスを実施した。つまり、本プロセスバリデーションにおいて4つの検量モデルについてそれぞれ3回のメンテナンスを実施した。全てのメンテナンスにおいて、モデル検証の結果がその基準に適合したため、Fig. 16 に示すメンテナンス開始、モデル検証、基準適合 (Yes)、及びメンテナンス終了の流れが実施された。メンテナンス開始から終了までの全ての手順は問題なく完了し、製造技術担当者や研究所への相談は発生しなかった。

また、検量モデルメンテナンス SOP を改訂した後に生産工場の担当者から意見を収集した結果、「SOP 改訂以前はモデルメンテナンスを実施する度に作業の流れを確認し、その対応方法を協議することが必要であった。SOP 更新以後はモデルメンテナンスに関する手順がより明確化された。これによって、GMP 管理下での適切性を維持しつつ、より効率的にモデルメンテナンスを実施できるようになった。作業負荷が著しく軽減されたため、担当者としては非常に有り難い」、「RACI 表を作製したことで、それぞれの作業で文書を回覧または情報を展開すべき担当者が明確化された。これによって情報伝達が遅延する危険性が低減された」、及び「このような手順や役割を明確化し、共有できる形式に残しておくことは適切に運用する上で非常に重要である。このような情報は、日常的な作業だけではなく、担当者間で業務を引き継ぐ際に

も役立つだろう」との意見を得た。

以上の結果から、本メンテナンスは適切かつ効率的に運用化されたと判断した。

### 3.4 小括

本章では、検量モデルをメンテナンスする具体的な枠組み及び商用生産工場における検量モデルメンテナンスを知識統合に基づいて運用化することを提案した。

検量モデルメンテナンスの重要性及び特定の評価手法が概念としては報告されているものの<sup>30), 31), 32)</sup>、その具体的な枠組み及び運用化に向けた方法論はこれまで報告されていない。そこで、まず各種のガイドライン<sup>31), 32)</sup>及びUSP<sup>33)</sup>を参考に検量モデルメンテナンスの具体的な枠組みを新規に構築した。この枠組みをGMP管理下で運用する上では、詳細な作業手順及び役割を明確化し、文書化することが求められる。しかし、複数の部署が関与する新規の枠組みに対して、あらゆる事態に対応する詳細な作業手順及び役割を予め明確化することは困難である。そこで、検量モデルメンテナンスを適切かつ効率的に運用化する方法論として、各部署が有する知識を統合することを検討した。知識を統合する手段として、業務プロセスモデル化技術であるIDEFO及び役割分担表RACIを用いた。

この知識統合によって、複数の部署が関与する新規の枠組みに対して、想定されたあらゆる事態に対応する詳細な作業手順及びGMP管理下でメンテナンス業務を運用するための必要十分な役割を明確化することができた。その結果、GMP管理下における運用の適切性を維持した上でより効率的に検量モデルメンテナンスを運用することが可能となった。なお、検量モデルメンテナンスにおいてモデル更新を実施する場合には検量モデルに使用する波数を改めて選択することになる。この波数選択において、第一章及び第二章で提案した波数選択法を用いることで高い推定精度を有する検量モデルを効率的に構築することができる。知識統合の議論の過程で、継続的なメンテナンス業務で得られる経験的な知識を製造プロセスの継続的な改善に活用する必要があることが判明し、知識管理の水準を向上させる契機となった。以上を踏まえて、検量モデルメンテナンスを商用生産工場において運用化した。運用手順を検証するために製造工程のプロセスバリデーション時に4つの検量モデルについて各3回のメンテナンスを実施した。その結果、全てのメンテナンスが問題なく完了したことから、検量モデルメンテナンスを適切に運用化できたと判断した。従って、当該工場における日常生産において本メンテナンスが継続的に使用されている。

これらの結果から、業務プロセスモデル化に基づいて知識を統合することが、検量モデルメンテナンスを適切かつ効率的に運用化するための方法論として有用であることが示された。

## 総括

製薬産業では、製品品質をより高度に保証するためのパラダイムとして QbD 及び PAT が注目されている。PAT の手段として NIRS が一般的に用いられている。NIRS に基づいて製剤の品質特性を推定するためには NIR スペクトルと推定対象である品質特性との関係を表す検量モデルが必要となる。検量モデルに関する最も重要な性能指標は推定精度であるが、特に波数選択はその推定精度に大きな影響を与える。また、製品のライフサイクルを通じて検量モデルの推定精度を恒常的に保証するために、検量モデルをメンテナンスすること、つまり継続的に推定精度を検証し、必要に応じて検量モデルを更新することが規制当局から求められている。

本論文では、NIRS に基づく検量モデルの推定精度を向上させるための効率的な波数選択法を開発するとともに、検量モデルメンテナンスを商用生産工場において適切に運用化する手法を提案した。

第一章では、化学物質の含量が変動するとその NIR 吸収帯に対応する特定の吸収ピークに影響を与えることに着目し、新規スペクトル分割法 SFD 及びその SFD に基づく波数選択法を提案した。従来の波数選択法である iPLS はスペクトル全体を等幅のスペクトル領域に分割し、そのスペクトル領域単位で波数を選択する。しかし、等幅のスペクトル領域では推定精度を向上させる上で有用な波数を柔軟に選択することができない。一方、提案するスペクトル分割法 SFD はスペクトル全体を各吸収ピークに対応するスペクトル領域に分割することを意図したものである。吸収ピークの境界ではスペクトル強度の変動が極小となると考えられる。そこで SFD は、まずキャリブレーションセットにおける各波数のスペクトル強度の標準偏差を算出し、それをスペクトル変動特性とする。次に、そのスペクトル変動特性の極小点においてスペクトル全体を複数のスペクトル領域に分割する。これによって、推定精度を向上させる上で有用な波数をより適切に選択できることが期待される。SFD に基づく波数選択法 SFD-PLS は、各スペクトル領域に含まれる全てのスペクトル強度を用いて PLS モデルを構築し、その推定精度が高い順にスペクトル領域を選択する。事例研究として、提案する SFD-PLS を顆粒中の水分含量及び薬物（化合物 X）含量を推定する 2 種類の検量モデル構築に適用した。SFD-PLS は従来法である PLS-beta、VIP、及び iPLS よりも高い推定精度を小さな計算負荷で達成した。この結果は、SFD-PLS が推定対象とする品質特性に関連する波数領域を最も正確にかつ効率的に選択したことを示唆している。さらに SFD-PLS に移動平均法を組合せた MASFD-PLS は推定精度を維持したまま分割するスペクトル領域の数を低減させた。以上の結果から、SFD-PLS 及び MASFD-PLS は従来法よりも推定精度及び計算負荷の観点で優れる有望な波数選択法と結論した。

第二章では、化学物質が一般に複数の NIR 吸収帯を有しており、複数のスペクトル領域におけるスペクトル強度が類似した挙動を示すという NIR スペクトルの特徴に着目し、スペクトル強度の相関性を利用した新規のクラスタリングに基づく波数選

択法を提案した。第一章では、小さな計算負荷で推定精度を改善する波数選択法として、スペクトル領域に基づく手法 **SFD-PLS** を提案した。**SFD-PLS** は各吸収ピークに対応するスペクトル領域を作製することを意図したものであり、従来法である **iPLS** よりも小さな計算負荷で高い推定精度を達成した。しかし、そのスペクトル領域の数は膨大であり、より効率的に高い推定精度を得る観点で改善の余地があった。そこで第二章では、各変動因子に由来する複数の吸収ピークに対応するスペクトル領域グループを作製することを意図した波数選択法 **SFD-NCSC-PLS** を新規に開発した。**SFD-NCSC-PLS** は、まず **NCSC** を用いて **SFD** に基づくスペクトル領域を数個のスペクトル領域グループにクラスタリングする。そして、各スペクトル領域グループに含まれる全てのスペクトル強度を用いて検量モデルを構築し、その検量モデルの推定精度が高い順にスペクトル領域グループを選択する。事例研究として、提案する **SFD-NCSC-PLS** を顆粒中の水分含量及び薬物（化合物 X）含量を推定する 2 種類の検量モデル構築に適用した。提案する **SFD-NCSC-PLS** は従来法である **iPLS** 及び **SCMWPLS** 並びに **SFD** または **NCSC** の一方のみを用いた手法よりも高い推定精度を達成した。この結果は、**SFD-NCSC-PLS** が推定対象である品質特性に関連する波数領域を最も正確に選択したことを示唆している。品質特性を推定する上で有用なスペクトル領域は様々な幅を有する。**iPLS** は等幅のスペクトル領域を作製するため、スペクトル領域を柔軟に選択することができない。**SCMWPLS** は **MWPLS** 及び **CSMWPLS** を組合せた手法であり、作製するスペクトル領域の幅を調節可能である。しかし、**MWPLS** は移動枠を用いて各波数点を近似的に評価し、その評価結果に基づいてスペクトル領域を作製する。さらに **CSMWPLS** は **MWPLS** に基づく各スペクトル領域から唯一つの副スペクトル領域を抽出する。ただし **MWPLS** に基づく各スペクトル領域には推定精度を向上させるために有用な副スペクトル領域が複数存在する可能性がある。こうした **SCMWPLS** のアルゴリズムでは、適切なスペクトル領域を選択し、推定精度を向上させることができない可能性がある。一方、**SFD-NCSC-PLS** は、各吸収ピークに対応するスペクトル領域を作製し、そのスペクトル領域をその領域面積の相関関係に基づいてクラスタリングすることで、ある変動因子に由来する複数の吸収ピークに対応するスペクトル領域グループを作製することを意図したものである。従って、**SFD-NCSC-PLS** は他の手法よりも正確に複数の吸収ピークに対応するスペクトル領域グループを選択することができたと考えられる。波数選択法を実用化する上で実際的な課題の 1 つは計算負荷であるが、**SFD-NCSC-PLS** は **SFD** または **NCSC** の一方のみを用いた手法よりも著しく小さい計算負荷で波数選択を完了した。以上の結果から、**SFD-NCSC-PLS** は第一章で提案した **SFD-PLS** よりもさらに小さな計算負荷で高い推定精度を達成する有望な波数選択法である。また、第一章及び第二章で提案した波数選択法は、異なる 2 つの推定対象に対して同様の結果が得られたことから、広範な推定対象に対してこれらの提案法を適用できると考えられる。本検討では顆粒のみを測定対象としたが、一般に **NIR** 吸収帯に応じて吸収ピークが生じ、その吸収ピークが

重なり合う等の NIR スペクトルの特徴は、測定対象に依らず同様である。従って、顆粒以外の様々な測定対象に対してもこれらの提案法を適用できるだろう。これらの提案法に基づく波数選択を実行するプログラムを作成しておけば、そのプログラムを実行することで自動的に波数選択を実施できる。従って、波数選択を実施する担当者に依らず同一の結果を得られる。また、推定対象や測定対象に依らず、同一のプログラムを用いて波数選択を実施できる。以上から、これらの提案法は推定精度及び計算負荷の観点で優れているだけでなく、波数選択結果が個人の能力に依存しない波数選択法であり、多様な推定対象及び測定対象に適用できると結論した。

第三章では、検量モデルをメンテナンスする具体的な枠組み及びそのメンテナンスを知識統合に基づいて運用化する方法論を提案した。検量モデルメンテナンスの重要性及び特定の評価手法が概念としては報告されているものの<sup>30), 31), 32)</sup>、その具体的な枠組み及び運用化に向けた方法論はこれまで報告されていない。そこで、まず各種のガイドライン<sup>31), 32)</sup>及び USP<sup>33)</sup>を参考に検量モデルメンテナンスの具体的な枠組みを新規に構築した。この枠組みを GMP 管理下で運用する上では、詳細な作業手順及び役割を明確化し、文書化することが求められる。しかし、複数の部署が関与する新規の枠組みに対して、あらゆる事態に対応する詳細な作業手順及び役割を予め明確化することは困難である。そこで、検量モデルメンテナンスを適切かつ効率的に運用化する方法論として、各部署が有する知識を統合することを検討した。知識を統合する手段として、業務プロセスモデル化技術である IDEF0 及び役割分担表 RACI を用いた。この知識統合によって、複数の部署が関与する新規の枠組みに対して、想定されたあらゆる事態に対応する詳細な作業手順及び GMP 管理下でメンテナンス業務を運用するための必要十分な役割を明確化することができた。その結果、GMP 管理下における運用の適切性を維持した上で、より効率的に検量モデルメンテナンスを運用することが可能となった。なお、検量モデルメンテナンスにおいてモデル更新を実施する場合には検量モデルに使用する波数を改めて選択することになる。この波数選択において、第一章及び第二章で提案した波数選択法を用いることで高い推定精度を有する検量モデルを効率的に構築することができる。知識統合の議論の過程で、継続的なメンテナンス業務で得られる経験的な知識を製造プロセスの継続的な改善に活用する必要があることが判明し、知識管理の水準を向上させる契機となった。以上を踏まえて、確立した検量モデルメンテナンスを商用生産工場にて運用化した。運用手順を検証するために製造工程のプロセスバリデーション時に4つの検量モデルについて各3回のメンテナンスを実施した。その結果、全てのメンテナンスが問題なく完了したことから、検量モデルメンテナンスを適切に運用化できたと判断した。従って、本工場における日常生産において本メンテナンスが継続的に使用されている。これらの結果から、業務プロセスモデル化に基づいて知識を統合することが、検量モデルメンテナンスを適切かつ効率的に運用化するために有用であると結論した。

製薬産業では製品品質及び生産効率の向上を目的として PAT に基づく迅速な品



質測定が実施されている。しかし、PAT が導入されるのは一部の限られた製品及び製造工程に対してのみである。この原因の一つとして PAT を製造工程に導入するために膨大な工数が必要とされることが挙げられる。特に PAT として一般的に用いられる NIRS に基づく検量モデルを開発し、検量モデルを継続的にメンテナンスするための工数が大きな負担となり得る。本研究で提案した効率的な波数選択法及び知識統合に基づく検量モデルメンテナンス運用化は、PAT を効率的に導入及び運用することに貢献するものである。提案した波数選択法は効率的であるだけでなく、検量モデルの推定精度を向上させることができるため、製品品質をより精密に測定及び管理することにも役立つことが期待される。さらに、本波数選択法は測定対象や推定対象に依らず適用できる。また本波数選択法は自動的に波数選択を実施できることから、波数選択結果が個人の能力に依存しない波数選択法である。

また、業務プロセスモデル化に基づく知識統合は、検量モデルメンテナンスに限らず製品品質を向上させるために今後導入される様々な新規技術を適切かつ効率的に運用化するための有用な手段となることが期待される。

## 謝辞

本研究の共同研究者として終始懇切なる御指導を賜りました京都大学大学院情報学研究科システム科学専攻ヒューマンシステム論分野 加納学教授、東京大学大学院工学系研究科化学システム工学専攻 杉山和弘准教授、及び第一三共株式会社製剤技術研究所 渡部知行氏に深謝致します。

本研究の遂行にあたり、多大な御協力を賜りました第一三共株式会社製薬技術本部製剤技術研究所 中川弘司氏及び田邊秀章氏、並びに京都大学大学院情報学研究科システム科学専攻ヒューマンシステム論分野 藤原幸一助教に深謝いたします。

本研究に対し、御配慮と御便宜を賜りました第一三共株式会社製剤技術研究所南秀実所長及び第一三共プロファーマ株式会社 脇山尚樹工場長に心より感謝申し上げます。

本論文の予備審査において、貴重なご助言及び懇切なる御指導を賜りました東京薬科大学薬学部医療衛生薬学科分析化学教室 袴田秀樹教授、東京薬科大学薬学部医療薬学科情報教育研究センター土橋朗教授、東京薬科大学薬学部医療薬物薬学科薬事関係法規教室 益山光一教授に心より感謝申し上げます。

最後に、本論文をまとめるにあたり、終始懇切なる御指導及び御鞭撻を賜りました東京薬科大学薬学部医療薬物薬学科製剤設計学教室 瀬田康生教授に心より感謝申し上げます。

## 引用文献

- 1) FDA, 2004. Guidance for industry, PAT - A framework for innovative pharmaceutical development, manufacturing, and quality assurance. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm070305.pdf>
- 2) EMA, 2006. Quality risk management (International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use Q9). [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Scientific\\_guideline/2009/09/WC500002873.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002873.pdf)
- 3) EMA, 2008. Note for guidance on pharmaceutical quality system (International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use Q10). [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Scientific\\_guideline/2009/09/WC500002871.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002871.pdf)
- 4) EMA, 2009. Note for guidance on pharmaceutical development. [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Scientific\\_guideline/2009/09/WC500002872.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002872.pdf)
- 5) ICH, ICH Harmonised Tripartite Guideline-Pharmaceutical Quality System Q10. 2005. [http://www.ich.org/fileadmin/Public\\_Web\\_Site/ICH\\_Products/Guidelines/Quality/Q10/Step4/Q10\\_Guideline.pdf](http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q10/Step4/Q10_Guideline.pdf).
- 6) ICH, ICH Harmonised Tripartite Guideline-Quality Risk Management Q9. 2005. [http://www.ich.org/fileadmin/Public\\_Web\\_Site/ICH\\_Products/Guidelines/Quality/Q9/Step4/Q9\\_Guideline.pdf](http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q9/Step4/Q9_Guideline.pdf).
- 7) ICH, ICH Harmonised Tripartite Guideline-Pharmaceutical Development Q8 (R2). 2009. [http://www.ich.org/fileadmin/Public\\_Web\\_Site/ICH\\_Products/Guidelines/Quality/Q8\\_R1/Step4/Q8\\_R2\\_Guideline.pdf](http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q8_R1/Step4/Q8_R2_Guideline.pdf).
- 8) Kourti, T., Davis, B., 2012. The business benefits of quality by design (QbD). *Pharm. Eng.* 32, 4.
- 9) Cogdill, R.P., Knight, T.P., Anderson, C.A., Drennen III, J.K. 2007. The financial returns on investments in process analytical technology and lean manufacturing: benchmarks and case study. *J. Pharm. Innov.* 2, 38–50.
- 10) Reich, G., 2005. Near-infrared spectroscopy and imaging: basic principles and pharmaceutical applications. *Adv. Drug Deliv. Rev.* 57, 1109–1143.
- 11) Roggo, Y., Chalus, P., Maurer, L., Lema-Martinez, C., Edmond, A., Jent, N., 2007. A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies. *J. Pharm. Biomed. Anal.* 44, 683–700.
- 12) Rantanen, J., Antikainen, O., Mannermaa, J.P., Yliruusi, J., 2000. Use of the near-infrared reflectance method for measurement of moisture content during granulation. *Pharm. Dev. Technol.* 5, 209–217.
- 13) Frake, P., Greenhalgh, D., Grierson, S.M., Hempenstall, J.M., Rudd, D.R., 1997. Process control and end-point determination of a fluid bed granulation by application of near infra-red spectroscopy. *Int. J. Pharm.* 151, 75–80.
- 14) Wu, H., Khan, M.A., 2009. Quality-by-design (QbD): An integrated approach for evaluation of powder blending process kinetics and determination of powder blending endpoint. *J. Pharm. Sci.* 98, 2784–2798.
- 15) Nakagawa, H., Kano, M., Hasebe, S., Suzuki, T., Wakiyama, N., 2013. Real-time monitoring of lubrication properties of magnesium stearate using NIR spectrometer and thermal effusivity sensor. *Int. J. Pharm.* 441, 402–413.

- 16) Kirsch, J.D., Drennen, J.K., 1995. Determination of film-coated tablet parameters by near-infrared spectroscopy. *J. Pharm. Biomed. Anal.* 13, 1273–1281.
- 17) Kirsch, J.D., Drennen, J.K., 1996. Near-infrared spectroscopic monitoring of the film coating process. *Pharm. Res.* 13, 234–237.
- 18) Porfire, A., Rus, L., Vonica, A.L., Tomuta, I., 2012. High-throughput NIR-chemometric methods for determination of drug content and pharmaceutical properties of indapamide powder blends for tableting. *J. Pharm. Biomed. Anal.* 70, 301–309.
- 19) Chalus, P., Roggo, Y., Walter, S., Ulmschneider, M., 2005. Near-infrared determination of active substance content in intact low-dosage tablets. *Talanta.* 66, 1294–1302.
- 20) Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics, *Chemom. Intell. Lab. Syst.* 58, 109–130.
- 21) Karande, A.D., Liew, C.V., Heng, P.W.S., 2010. Calibration sampling paradox in near infrared spectroscopy: A case study of multi-component powder blend. *Int. J. Pharm.* 395, 91–97.
- 22) Jia, Y.H., Liu, X.P., Feng, Y.C., Hu, C.Q., 2011. A training set selection strategy for a universal near-infrared quantitative model. *AAPS Pharm. Sci. Tech.* 12, 738–745.
- 23) Shi, Z., Hermiller, J.G., Gunter, T.Z., Zhang, X., Reed, D.E., 2012. A novel sample selection strategy by near-infrared spectroscopy-based high throughput tablet tester for content uniformity in early-phase pharmaceutical product development. *J. Pharm. Sci.* 101, 2502–2511.
- 24) Savitzky, A., Golay, M.J., 1964. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* 36, 1627–1639.
- 25) Barnes, R., Dhanoa, M., Lister, S.J., 1989. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Appl. Spectrosc.* 43, 772–777.
- 26) Xu, L., Zhou, Y.P., Tang, L.J., Wu, H.L., Jiang, J.H., Shen, G.L., Yu, R.Q., 2008. Ensemble preprocessing of near-infrared (NIR) spectra for multivariate calibration. *Anal. Chim. Acta.* 616, 138–143.
- 27) Xiaobo, Z., Jiewen, Z., Povey, M.J., Holmes, M., Hanpin, M., 2010. Variables selection methods in near-infrared spectroscopy. *Analytica Chimica Acta.* 667, 14–32.
- 28) Nadler, B., and Coifman, R.R. 2005. The prediction error in CLS and PLS: the importance of feature selection prior to multivariate calibration. *J. Chemom.* 19, 107–118.
- 29) Igne, B., Shi, Z., Drennen, J.K., Anderson, C.A., 2014. Effects and detection of raw material variability on the performance of near-infrared calibration models for pharmaceutical products. *J. Pharm. Sci.* 103, 545–556.
- 30) Kano, M., Ogawa, M., 2010. The state of the art in chemical process control in Japan: Good practice and questionnaire survey. *J. Proc. Cont.* 20, 969–982.
- 31) FDA, Code of Federal Regulations (CFR) Title 21. 2013. <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=211&showFR=1>.
- 32) EMA, Guideline on the use of near infrared spectroscopy by the pharmaceutical industry and the data requirements for new submissions and variations. 2014. [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Scientific\\_guideline/2014/06/WC500167967.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/06/WC500167967.pdf)
- 33) The United States Pharmacopeial Convention. United States Pharmacopeia. 37th revision, NF 32; 2014. pp. 953–959.
- 34) Kim, S., Kano, M., Nakagawa, H., Hasebe, S., 2011. Estimation of active pharmaceutical ingredients content using locally weighted partial least squares and statistical wavelength

- selection. *Int. J. Pharm.* 421, 269–274.
- 35) Chong, I.G., Jun, C.H., 2005. Performance of some variable selection methods when multicollinearity is present. *Chemom. Intell. Lab. Syst.* 78, 103–112.
  - 36) Nørgaard, A.S.L., Wagner, J., Nielsen, J.P., Munck, L., Engelsen, S.B., 2000. Interval Partial Least-Squares Regression (iPLS): A Comparative Chemometric Study with an Example from Near-Infrared Spectroscopy. *Appl. Spectrosc.* 54, 413–419.
  - 37) Jiang, J.H., Berry, R.J., Siesler, H.W., Ozaki, Y., 2002. Wavelength interval selection in multicomponent spectral analysis by moving window partial least-squares regression with applications to mid-infrared and near-infrared spectroscopic data. *Anal. Chem.* 74, 3555–3565.
  - 38) Du, Y.P., Liang, Y.Z., Jiang, J.H., Berry, R.J., Ozaki, Y., 2004. Spectral regions selection to improve prediction ability of PLS models by changeable size moving window partial least squares and searching combination moving window partial least squares. *Analytica Chimica Acta.* 501, 183–191.
  - 39) Arakawa, M., Yamashita, Y., Funatsu, K., 2011. Genetic algorithm-based wavelength selection method for spectral calibration. *J. Chemom.* 25, 10–19.
  - 40) Leardi, R., Nørgaard, L., 2004. Sequential application of backward interval partial least squares and genetic algorithms for the selection of relevant spectral regions. *J. Chemom.* 18, 486–497.
  - 41) ICH, Quality implementation working group Points to Consider (R2) ICH-endorsed guide for ICH Q8/Q9/Q10 implementation. 2011. [http://www.ich.org/fileadmin/Public\\_Web\\_Site/ICH\\_Products/Guidelines/Quality/Q8\\_Q9\\_Q10\\_QAs/PtC/Quality\\_IWG\\_PtCR2\\_6dec2011.pdf](http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q8_Q9_Q10_QAs/PtC/Quality_IWG_PtCR2_6dec2011.pdf)
  - 42) European Directorate for the Quality of Medicines & HealthCare. *European Pharmacopoeia*. 8th Edition; 2014. pp. 62–67.
  - 43) Yamatera, H., Fitzpatrick, B., Gordon, G., 1964. Near infrared spectra of water and aqueous solutions. *J. Mol. Spectrosc.* 14, 268–278.
  - 44) Fujiwara, K., Sawada, H., Kano, M., 2012. Input variable selection for PLS modeling using nearest correlation spectral clustering. *Chemom. Intell. Lab. Syst.*, 118, 109–119.
  - 45) Fujiwara, K., Kano, M., Hasebe, S., 2010. Development of correlation-based clustering method and its application to software sensing. *Chemom. Intell. Lab. Syst.* 101, 130–138.
  - 46) Fujiwara, K., Kano, M., Hasebe, S., 2011. Correlation-based spectral clustering for flexible process monitoring. *J. Process Control* 21, 1438–1448.
  - 47) Kano, M., Fujiwara, K., 2013. Virtual sensing technology in process industries: trends and challenges revealed by recent industrial applications. *J. Chem. Eng. Japan* 46, 1–17.
  - 48) Ng, A.Y., Jordan, M.I., Weiss, Y., 2002. On spectral clustering: Analysis and an algorithm. *Adv. Neural Inf. Process. Syst.* 2, 849–856.
  - 49) Ross, D.T., 1985. Application and Extensions of SADT. *Computer.* 18, 25–34.
  - 50) National Institute of Standards and Technology (NIST). 1993. Integration Definition for Function Modeling (IDEF0), Federal Information Processing Standards Publication No. 183.
  - 51) Kawai, H., Seki, H., Fuchino, T., Naka, Y., 2012. Pharmaceutical engineering strategy for quality informatics on the IDEF0 business process model. *J. Pharm. Innov.* 7, 195–204.
  - 52) Jacka, J.M., Keller, P.J., 2009. *Business process mapping: improving customer satisfaction*. John Wiley & Sons.
  - 53) Sugiyama, H., Schmidt, R., 2014. Business process model of continuous improvement in pharmaceutical manufacturing. *Kagaku Kogaku Ronbunshu.* 40, 201–210.
  - 54) Fuchino, T., Shimada, Y., Kitajima, T., Takeda, K., Batres, R., Naka, Y., 2011. A business

- process model for process design that incorporates independent protection layer considerations. *Comput. Aided Chem. Eng.* 29, 326–330.
- 55) Sugiyama, H., Hirao, M., Fischer, U., Hungerbühler, K., 2008. Activity modeling for integrating environmental, health and safety (EHS) consideration as a new element in industrial chemical process design. *J. Chem. Eng. Japan.* 41, 884–997.
  - 56) Blokdijs, G., Menken, I., 2008. *The Service Level Agreement SLA Guide-SLA Book, Templates for Service Level Management and Service Level Agreement Forms. Fast and Easy Way to Write Your SLA.* Emereo Pty Limited.
  - 57) Hale, J., 2003. *Performance-based management: What every manager should do to get results.* John Wiley & Sons.
  - 58) Jackson, J., Mudholkar, G., 1979. Control procedures for residuals associated with principal component analysis. *Technometrics.* 21, 341–349.
  - 59) Kano, M., Hasebe, S., Hashimoto, I., Ohno, H., 2002. Statistical process monitoring based on dissimilarity of process data. *AIChE J.* 48, 1231–1240.
  - 60) Kresta, J.V., MacGregor, J.F., Marlin, T.E., 1991. Multivariate statistical monitoring of process operating performance. *Can. J. Chem. Eng.* 69, 35–47.
  - 61) Western E., 1956. *Statistical Quality Control Handbook.* Western Electric Co., Inc., 25–28.
  - 62) Nakagawa, H., Kano, M., Hasebe, S., Miyano, T., Watanabe, T., Wakiyama, N., 2014. Verification of model development technique for NIR-based real-time monitoring of ingredient concentration during blending. *Int. J. Pharm.* 47, 264–275.