



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0041564

(51)^{2020.01} G01N 21/27; G16C 60/00; G16C 20/00; (13) B
G16C 20/30; G01N 21/35; G01N 21/359

(21) 1-2022-02966 (22) 05/10/2020
(86) PCT/EP2020/077769 05/10/2020 (87) WO2021/073908 22/04/2021
(30) 19203750.5 17/10/2019 EP
(45) 25/11/2024 440 (43) 25/10/2022 415
(73) EVONIK OPERATIONS GMBH (DE)
Rellinghauser Strasse 1-11, 45128 Essen, Germany
(72) REIMANN, Ingolf (DE); REISING, Joachim (DE).
(74) Công ty Luật TNHH Phạm và Liên danh (PHAM & ASSOCIATES)

(54) PHƯƠNG PHÁP VÀ HỆ THỐNG DỰ ĐOÁN GIÁ TRỊ THUỘC TÍNH QUAN TÂM CỦA NGUYÊN LIỆU

(57) Sáng chế đề cập đến phương pháp dự đoán giá trị thuộc tính quan tâm của nguyên liệu được thực hiện trên máy tính, bao gồm các bước sau: a) xác định quần thể phổ hồng ngoại của các mẫu, trong đó các phổ này tạo ra $m \times n$ ma trận dữ liệu đầu vào X , với m là số lượng mẫu trong các hàng và n là các điểm dữ liệu trong các cột, b) loại bỏ các giá trị ngoại lệ phổ ra khỏi quần thể phổ thu được ở bước a), bao gồm các bước sau: b1) thu nhận các thành phần chính bằng cách xử lý ma trận X bằng phân tích thành phần chính, b2) tạo ra ma trận chéo Σ , chứa các giá trị suy biến σ_m của ma trận X , và ma trận hệ số tải V , từ ma trận dữ liệu đầu vào X , b3) tính toán điểm số x_m đối với mỗi phổ bằng cách nhân mỗi điểm dữ liệu của ma trận dữ liệu đầu vào X với các hệ số tải đối với mỗi thành phần thu được ở bước b2), tạo ra giá trị trung bình của mỗi cột của ma trận X để tạo ra các giá trị $B_{0,m}$, và tính toán chỉ số chấm điểm si bằng công thức $si = \sum_i \frac{(x_m - B_{0,m})}{\sigma_m}$, b4) xác định số lượng của các thành phần N_C có các giá trị giềng dẫn đến sự hội tụ trong hồi quy của X trên các điểm số ít nhất bằng 99%, và tính toán giá trị giới hạn đo khoảng cách đối với mọi phổ thu được ở bước a), bằng công thức $D_i = N_C \times \sqrt{\frac{(si^2 \times N_C)}{(N_C - 1)}}$, b5) tính toán giá trị trung bình đối với toàn bộ các điểm số của mỗi thành phần chính của mỗi phổ thu được ở bước a) và tính toán giá trị đo khoảng cách giữa giá trị trung bình này và mỗi điểm số của mỗi thành phần chính, b6) xem xét phổ mẫu là giá trị ngoại lệ phổ khi giá trị của giá trị đo khoảng cách đối với điểm số của thành phần chính thu được ở bước b5) lớn hơn giá trị giới hạn đo khoảng cách thu được ở bước b4), b7) loại bỏ giá trị ngoại lệ phổ thu được ở bước b6) ra khỏi quần thể phổ thu được ở bước a) để thu được quần thể phổ được làm sạch, c) tạo ra hàm dự đoán trên quần thể phổ được làm sạch thu được ở bước b7), d) xác định phổ hồng ngoại của mẫu có nguồn gốc và/hoặc thành phần chưa biết hoặc có cùng nguồn gốc và/hoặc thành phần như mẫu ở bước a), và e) dự đoán giá trị thuộc tính quan tâm từ phổ thu được ở bước d) bằng hàm dự đoán thu được ở bước c). Sáng chế cũng đề cập đến hệ thống dự đoán giá trị thuộc tính quan tâm của nguyên liệu.

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp dự đoán giá trị thuộc tính quan tâm của nguyên liệu được thực hiện trên máy tính và hệ thống để thực hiện phương pháp này bao gồm bộ xử lý.

Tình trạng kỹ thuật của sáng chế

Phổ hồng ngoại gần là công cụ hữu ích để dự đoán giá trị thuộc tính quan tâm của nguyên liệu, cụ thể khi cách xa phòng thí nghiệm phân tích trong đó phân tích, định tính và/hoặc định lượng, thường sẽ được thực hiện. Cụ thể, các sản phẩm nông nghiệp, như nguyên liệu thức ăn chăn nuôi và/hoặc thức ăn chăn nuôi, có thể được phân tích về sự có mặt và nồng độ của các thành phần cụ thể trước và sau các giai đoạn xử lý như bước rang và xử lý bằng áp suất, trong hoặc sau khi bảo quản và sau khi chế biến thức ăn chăn nuôi chứa chúng. Tuy nhiên, cần có hàm hiệu chuẩn thích hợp để tạo ra các phương pháp dự đoán chính xác và tin cậy bằng phổ hồng ngoại gần.

Hàm hiệu chuẩn này thường được tạo ra bằng phân tích đa biến. Điều này cho phép xem xét thích hợp các mối tương quan để ước lượng thành phần của các hỗn hợp phức tạp với mức bù cho các tín hiệu nhiễu từ các tín hiệu nền. Việc dự đoán giá trị thuộc tính quan tâm bằng phổ hồng ngoại gần và hàm hiệu chuẩn thể hiện quy trình hai bước. Trong bước thứ nhất, mô hình hiệu chuẩn được xây dựng bằng cách sử dụng tập dữ liệu thu được bằng các phép đo gián tiếp, ví dụ cường độ tín hiệu quang học, và các phép đo trực tiếp, ví dụ nồng độ của các chất phân tích, trong một số tình huống mở rộng độ đa dạng của các điều kiện thiết bị và chất phân tích khác nhau. Biểu thức chung đối với mối tương quan giữa các phép đo trực tiếp (ví dụ nồng độ của chất phân tích) và các phép đo gián tiếp (ví dụ cường độ tín hiệu quang học) là $y = f(x_1, x_2, \dots, x_n)$, trong đó y là giá trị thuộc tính quan tâm mong muốn cần dự đoán (ví dụ nồng độ của chất phân tích), f là một số hàm (mô hình), và $x_1, x_2, \dots, x_n$ là các đối số của mô hình, và cụ thể là các kết quả của các phép đo gián tiếp bất kỳ, ví dụ các phép đo quang học được biến đổi ở số lượng (cụ thể) của các bước sóng. Mục đích của bước thứ nhất này là phát triển hàm hữu dụng f phản ánh mối tương quan giữa các phép đo gián tiếp và giá trị thuộc tính quan tâm mong muốn cần dự

đoán. Trong bước thứ hai (dự đoán), hàm này được đánh giá bằng tập đo được của các phép đo gián tiếp (quang học) $x_1, x_2, \dots, x_n$ để thu được giá trị ước lượng của phép đo trực tiếp (ví dụ nồng độ của chất phân tích) ở một số thời điểm trong tương lai khi các phép đo quang học sẽ được thực hiện mà không có phép đo trực tiếp tương ứng.

Có một số lượng lớn tài liệu chuyên ngành liên quan đến việc tạo ra hàm hiệu chuẩn để dự đoán giá trị thuộc tính quan tâm của nguyên liệu bằng phổ hồng ngoại gần. Hầu hết các tài liệu đã biết này đề cập đến các khái niệm chung và cụ thể để tạo ra hàm hiệu chuẩn. Một số tài liệu chuyên ngành đề cập đến việc bù các tín hiệu nhiễu, ví dụ các tín hiệu nhiễu từ môi trường và các sai số đặc trưng thiết bị như các sai số phép đo và các nguồn phát xạ hóa già, trong việc tạo ra hàm hiệu chuẩn. Tuy nhiên, các tài liệu đã biết không xem xét việc xác định các giá trị ngoại lệ phổ, hướng chỉ cách thức phân bố chúng trong việc tạo ra hàm hiệu chuẩn nhưng rất có thể chúng sẽ xuất hiện trong quá trình tích lũy các phép đo gián tiếp (quang học). Do đó, rất có thể chúng sẽ làm giảm việc tạo ra hàm hiệu chuẩn. Giá trị ngoại lệ có thể là do độ biến thiên trong phép đo hoặc có thể là kết quả của sai số thiết bị; kết quả của sai số thiết bị đôi khi được loại trừ ra khỏi tập dữ liệu. Nhìn chung, và cụ thể trong thống kê, giá trị ngoại lệ được xem là điểm dữ liệu khác biệt đáng kể so với các giá trị quan sát khác, tuy nhiên để lại nhiều không gian cho việc diễn giải chủ quan và diễn giải sai. Mặt khác, việc bao gồm các điểm dữ liệu ở các vùng biên của tập dữ liệu cần thiết để hiệu chuẩn có ý nghĩa và thô, do đó không nên bỏ qua đơn giản chỉ vì chúng có vẻ dư. Điều này đã cho thấy rằng một, nếu không phải là vấn đề chính liên quan đến các giá trị ngoại lệ là sự phát hiện hoặc nhận dạng của chúng vì không có định nghĩa cứng nhắc của giá trị ngoại lệ. Do đó, cuối cùng đây vẫn là phép kiểm định chủ quan để xác định xem giá trị quan sát có là giá trị ngoại lệ hay không. Do thiếu định nghĩa được chấp nhận chung của giá trị ngoại lệ, nên có nhiều phương pháp khác nhau để phát hiện giá trị ngoại lệ. Một số phương pháp là phương pháp đồ thị như đồ thị xác suất chuẩn, một số phương pháp khác là phương pháp lai trên cơ sở mô hình như biểu đồ hộp, và việc lựa chọn phương pháp và cách thức phân bố giá trị ngoại lệ thường phụ thuộc vào từng trường hợp. Ngay cả khi mô hình phân bố chuẩn thích hợp với dữ liệu đang được phân tích, các giá trị ngoại lệ phải được mong đợi cho các cỡ mẫu lớn và không được tự động loại bỏ khi chúng có mặt. Ứng dụng tương ứng nên sử dụng thuật toán phân loại chính xác đối với các giá trị ngoại lệ cho dữ liệu mô hình với các giá trị ngoại lệ xuất hiện tự nhiên. Việc loại bỏ dữ liệu giá trị ngoại lệ là thực tiễn gây tranh cãi được nhiều nhà khoa học và giảng viên khoa học phản đối.

Việc loại bỏ dữ liệu giá trị ngoại lệ được chấp nhận nhiều hơn trong các lĩnh vực thực hành, trong đó mô hình cơ bản của quy trình và phân bố chuẩn của các sai số phép đo đã được biết rõ. Giá trị ngoại lệ thu được từ sai số đọc thiết bị có thể được loại trừ nhưng mong muốn là sai số đọc thiết bị ít nhất được kiểm chứng. Cụ thể, trong trường hợp của phổ hồng ngoại, đó là thách thức lớn để xác định các giá trị ngoại lệ chính xác nhất có thể và cũng rất hiệu quả, tốt nhất là cả hai cùng một lúc. Điều này thậm chí còn phù hợp hơn khi các phổ hồng ngoại tương ứng (các phép đo gián tiếp) và dữ liệu tham chiếu tương ứng (các phép đo trực tiếp) sẽ là cơ sở cho việc tạo ra hàm hiệu chuẩn thích hợp để dự đoán giá trị thuộc tính quan tâm của nguyên liệu bằng phổ hồng ngoại.

Do đó, cần có phương pháp dự đoán giá trị thuộc tính quan tâm của nguyên liệu bằng phổ hồng ngoại, cho phép xác định và loại bỏ tự động và đáng tin cậy các giá trị ngoại lệ phổ trong quá trình tạo ra hàm hiệu chuẩn.

Bản chất kỹ thuật của sáng chế

Do đó, mục đích của sáng chế là đề cập đến phương pháp dự đoán giá trị thuộc tính quan tâm của nguyên liệu được thực hiện trên máy tính, bao gồm các bước sau:

a) xác định quần thể phổ hồng ngoại của các mẫu, trong đó các phổ này tạo ra $m \times n$ ma trận dữ liệu đầu vào X , với m là số lượng mẫu trong các hàng và n là các điểm dữ liệu trong các cột,

b) loại bỏ các giá trị ngoại lệ phổ ra khỏi quần thể phổ thu được ở bước a), bao gồm các bước sau:

b1) thu nhận các thành phần chính bằng cách xử lý ma trận X bằng phân tích thành phần chính,

b2) tạo ra ma trận chéo Σ , chứa các giá trị suy biến σ_m của ma trận X , và ma trận hệ số tải V , từ ma trận dữ liệu đầu vào X ,

b3) tính toán điểm số x_m đối với mỗi phổ bằng cách nhân mỗi điểm dữ liệu của ma trận dữ liệu đầu vào X với các hệ số tải đối với mỗi thành phần thu được ở bước b2), tạo ra giá trị trung bình của mỗi cột của ma trận X để tạo ra các giá trị $B_{0,m}$, và tính toán chỉ số chấm điểm si bằng công thức

$$si = \sum_i \frac{(x_m - B_{0,m})}{\sigma_m}$$

b4) xác định số lượng của các thành phần N_c có các giá trị giềng dẫn đến sự hội tụ trong hồi quy của X trên các điểm số ít nhất bằng 99%, và tính toán giá trị giới hạn đo khoảng cách D_i đối với mọi phổ thu được ở bước a), bằng công thức

$$D_i = N_c \times \sqrt{\frac{(s_i^2 \times N_c)}{(N_c - 1)}}$$

b5) tính toán giá trị trung bình đối với toàn bộ các điểm số của mỗi thành phần chính của mỗi phổ thu được ở bước a) và tính toán giá trị đo khoảng cách giữa giá trị trung bình này và mỗi điểm số của mỗi thành phần chính,

b6) xem xét phổ mẫu là giá trị ngoại lệ phổ khi giá trị của giá trị đo khoảng cách đối với điểm số của thành phần chính thu được ở bước b5) lớn hơn giá trị giới hạn đo khoảng cách thu được ở bước b4),

b7) loại bỏ giá trị ngoại lệ phổ thu được ở bước b6) ra khỏi quần thể phổ thu được ở bước a) để thu được quần thể phổ được làm sạch,

c) tạo ra hàm dự đoán trên quần thể phổ được làm sạch thu được ở bước b7),

d) xác định phổ hồng ngoại của mẫu có nguồn gốc và/hoặc thành phần chưa biết hoặc có cùng nguồn gốc và/hoặc thành phần như mẫu ở bước a), và

e) dự đoán giá trị thuộc tính quan tâm từ phổ thu được ở bước d) bằng hàm dự đoán thu được ở bước c).

Mô tả chi tiết sáng chế

Quần thể phổ hồng ngoại của các mẫu được xác định ở bước a) xác định không gian dữ liệu n chiều, trong đó n là số lượng của các điểm dữ liệu trong các phổ và m là số lượng của các mẫu. Do đó, dữ liệu của các phổ hồng ngoại có thể được biểu diễn bằng ma trận dữ liệu đầu vào của loại $X_{m \times n}$ với m hàng đối với các mẫu và n cột đối với các điểm dữ liệu, thường cũng được viết dưới dạng ma trận $m \times n$ X . Lượng tuyệt đối của dữ liệu làm cho ma trận thu được khá phức tạp. Do đó, ma trận dữ liệu đầu vào được xử lý giảm dữ liệu, mà không làm thất thoát thông tin liên quan. Điều này thường được thực hiện trong phân tích thành phần chính (PCA). Đây là quy trình thống kê sử dụng phép biến đổi trực giao để biến đổi tập quan sát của các biến có thể tương quan (mỗi đại lượng của tập này nhận các giá trị khác nhau) thành tập giá trị của các biến không tương quan tuyến tính được gọi là các thành phần chính. Phép biến đổi này được xác định theo cách thức sao cho thành

phần chính thứ nhất có phương sai lớn nhất có thể, tức là nó dẫn đến độ biến thiết cao nhất có thể về dữ liệu. Mỗi thành phần khác trực giao với các thành phần trước đó và đứng thứ hai sau thành phần trước đó về phương sai cao nhất có thể. Các vector thu được (mỗi vector là tổ hợp tuyến tính của các biến và bao gồm n quan sát) là tập cơ sở trực giao không tương quan. Trong phân tích thành phần chính, ma trận dữ liệu đầu vào $X_{m \times n}$ được suy biến thành hai ma trận trực giao với nhau, ma trận $U_{m \times n}$ và ma trận $V_{n \times n}^T$ matrix. Về mặt toán học, bước này được mô tả bằng công thức

$$X_{m \times n} = U_{m \times n} \times V_{n \times n}^T$$

$V_{n \times n}^T$ là ma trận hệ số tải, bao gồm các hệ số tải, tức là các trục mới thu được từ phép biến đổi, và $U_{m \times n}$ cũng được gọi là ma trận điểm số, bao gồm các tọa độ mới. Các hệ số tải có thể được hiểu là các trọng số mà các biến ban đầu sẽ được nhân để tính toán các thành phần chính.

Theo sáng chế ma trận dữ liệu đầu vào X được suy biến thành tích số của một số ma trận bằng phép suy biến giá trị đơn, về mặt toán học được mô tả bằng

$$X_{m \times n} = U_{m \times m} \times \Sigma_{m \times n} \times V_{n \times n}^*$$

trong đó $X_{m \times n}$ là ma trận $m \times n$ của hạng k , $U_{m \times m}$ là ma trận đơn $m \times m$, $\Sigma_{m \times n}$ là ma trận thực $m \times n$ chứa các giá trị suy biến σ_m của ma trận X , và $V_{n \times n}^*$ là ma trận liên kế của ma trận đơn $n \times n$ V chứa các hệ số tải; với các số thực ma trận liên kế tương đương với ma trận chuyển vị $V_{n \times n}^T$.

Theo sáng chế quần thể phổ hồng ngoại không bị giới hạn bất kỳ liên quan đến số lượng của các phổ, đã xác định được rằng số lượng của các phổ là đủ để tạo ra hàm dự đoán có ý nghĩa. Do đó, số lượng của các phổ trong quần thể phổ có thể nằm trong khoảng từ 50 đến 10000, nằm trong khoảng từ 50 đến 5000, nằm trong khoảng từ 50 đến 2,500, nằm trong khoảng từ 50 đến 2000, nằm trong khoảng từ 50 đến 1500, nằm trong khoảng từ 50 đến 1000, nằm trong khoảng từ 100 đến 1000, nằm trong khoảng từ 50 đến 500, nằm trong khoảng từ 100 đến 500, nằm trong khoảng từ 50 đến 250, nằm trong khoảng từ 100 đến 250, hoặc nằm trong khoảng từ 50 đến 100.

Phổ hồng ngoại/các phổ hồng ngoại có thể được ghi ở các bước sóng nằm trong khoảng từ 400nm đến 2800nm bằng thiết bị quang phổ thích hợp bất kỳ hoạt động theo nguyên tắc đơn sắc hoặc nguyên tắc biến đổi Fourier. Tốt hơn nếu, phổ hồng ngoại được

ghi ở bước sóng nằm trong khoảng từ 1100nm đến 2500nm. Các bước sóng dễ dàng được biến đổi thành các số sóng tương ứng, do đó phổ hồng ngoại cũng có thể được ghi ở các số sóng tương ứng. Nhìn chung, phổ hồng ngoại yêu cầu sự có mặt của các nhóm chức hoặc các gốc chức, cho phép kích thích dao động phân tử trong nguyên liệu bằng cách chiếu xạ, sau đó tần số của ánh sáng chiếu qua được ghi bằng thiết bị quang phổ. Tuy nhiên, cũng có thể ghi các phổ hồng ngoại của nguyên liệu không có các nhóm chức hoặc các gốc chức bất kỳ, ví dụ các khoáng chất, được sử dụng riêng không cho phép các dao động phân tử bất kỳ. Tuy nhiên phổ hồng ngoại này yêu cầu sự có mặt đồng thời của chất có hoạt tính hồng ngoại, tương ứng với chất có hoạt tính không hồng ngoại, ví dụ trong sự hình thành của hợp chất phức tạp, và do đó phải chịu các thay đổi dao động trong kích thích các dao động phân tử. Do đó, bước so sánh phổ hồng ngoại thu được với phổ của chất có hoạt tính hồng ngoại (tinh khiết) trong sự vắng mặt của chất không có hoạt tính hồng ngoại bổ sung, cho phép rút ra kết luận về độ tương đồng và nồng độ của chất có không có hoạt tính hồng ngoại. Theo đó, phổ hồng ngoại cho phép dự đoán của giá trị thuộc tính quan tâm của nhiều nguyên liệu khác nhau. Tuy nhiên, mẫu sinh học như thức ăn chăn nuôi chứa nhiều hợp chất hữu cơ và vô cơ khác nhau, do đó biểu diễn ma trận khá phức tạp. Tuy nhiên, mọi chất sinh học tạo ra phổ hồng ngoại gần duy nhất, tương đương với dấu vân tay riêng biệt. Do đó, hai chất sinh học tạo ra các phổ giống hệt nhau có thể được giả định là có cùng thành phần hóa học và vật lý, do đó giống hệt nhau. Mặt khác, khi hai chất sinh học tạo ra các phổ khác nhau, có thể giả định rằng chúng khác nhau, về đặc tính vật lý hoặc đặc tính hóa học, hoặc cả hai đặc tính. Do các băng hấp thụ riêng lẻ và rất đặc trưng của chúng, nên các tín hiệu của các hợp chất hữu cơ và vô cơ và cường độ của chúng trong các phổ hồng ngoại có thể dễ dàng gán cho và tương quan với hợp chất hữu cơ cụ thể và nồng độ của nó trong mẫu có khối lượng đã biết. Do đó, phổ hồng ngoại cho phép dự đoán của giá trị thuộc tính quan tâm của nguyên liệu khá phức tạp, như độ tương đồng và nồng độ của các axit amin và protein khác nhau trong mẫu sinh học. Tuy nhiên, phổ hồng ngoại thậm chí cũng cho phép xác định an thông số lý thuyết trong các mẫu phức tạp, khi có thể xác định các xu hướng chung trong cùng loại mẫu do các thay đổi của các thông số lý thuyết này. Ví dụ, có thể gán và tương quan độ hấp thụ và cường độ của chúng trong phổ hồng ngoại của mẫu phức tạp của thức ăn chăn nuôi hoặc nguyên liệu thức ăn chăn nuôi cho các thông số lý thuyết, như hoạt tính chất ức chế trypsin, hoạt tính ureaza, độ hòa tan protein trong môi trường kiềm và chỉ số độ phân tán protein, và nồng độ của chúng. Trong bước tiếp theo,

thiết bị quang phổ hồng ngoại được sử dụng phải được hiệu chuẩn. Khi, cường độ hấp thụ ở các bước sóng hoặc các số sóng tương ứng đã được hiệu chỉnh thành công, tức là tức được gán và tương quan với các thông số quan tâm và giá trị của chúng, phổ hồng ngoại gần cho phép dự đoán chính xác giá trị thuộc tính quan tâm của nguyên liệu. Vì mục đích này, số lượng đầy đủ của các phổ hồng ngoại, khi cần 100, 200, 300, 400, 500 hoặc cao hơn, của nguyên liệu truy vấn được ghi, và cường độ hấp thụ ở các bước sóng hoặc các số sóng tương ứng được hiệu chỉnh phù hợp với các thông số tương ứng và các giá trị của chúng. Phổ hồng ngoại có thể được đo theo chế độ truyền qua hoặc phản xạ. Trong sử dụng nghiên cứu, chế độ đo phổ biến nhất là chế độ truyền qua nhưng để đo nguyên liệu không hòa tan hoặc dạng hạt, tốt hơn nếu đo theo chế độ phản xạ. Trong trường hợp đo theo chế độ phản xạ, độ phản xạ của ánh sáng trở lại từ mẫu được đo và độ chênh lệch giữa ánh sáng phát ra và phản xạ được tạo ra dưới dạng độ hấp thụ.

Nhìn chung, phương pháp theo sáng chế không bị giới hạn ở nguyên liệu có giá trị thuộc tính quan tâm cần dự đoán. Do đó, nguyên liệu này có thể là một chất duy nhất hoặc hỗn hợp của các chất khác nhau. Như được mô tả chi tiết nêu trên, nguyên liệu phải chứa ít nhất một chất có hoạt tính hồng ngoại để cho phép kích thích các dao động phân tử. Ngoài ra, nguyên liệu có thể có thành phần và/hoặc nguồn gốc bất kỳ. Các nguyên liệu thích hợp theo sáng chế là các hợp chất hữu cơ, như axit amin, peptit, protein, axit cacboxylic, axit béo, axit béo no và/hoặc không no, axit béo không no đa chức, hoặc hỗn hợp của chúng, hoặc nguyên liệu có nguồn gốc từ người, động vật hoặc thực vật, ví dụ phần thức ăn được tiêu hóa và hấp thụ, bộ phận cơ thể, bộ phận cơ thể động vật hoặc bộ phận cơ thể thực vật, hoặc các sản phẩm được sản xuất công nghiệp trên cơ sở các sản phẩm có nguồn gốc thực vật và/hoặc động vật, ví dụ thức ăn chăn nuôi, thức ăn dinh dưỡng cho động vật và/hoặc nguyên liệu thức ăn chăn nuôi, như hạt ngũ cốc chung chất chứa các chất hòa tan (DDGS), bột cá thủy phân hoặc bột lông vũ thủy phân, và các hợp chất vô cơ, như carbonat, phosphat, nitrat, và sulfat, hoặc hỗn hợp của chúng. Tốt hơn nếu, nguyên liệu có giá trị thuộc tính quan tâm cần dự đoán là axit amin, protein, peptit, axit cacboxylic, axit béo no, không no và/hoặc no một phần, axit béo không no đa chức, thức ăn dùng cho vật nuôi, thức ăn chăn nuôi, nguyên liệu thức ăn chăn nuôi, chất có nguồn gốc từ người, động vật hoặc thực vật hoặc hỗn hợp của chúng.

Nguyên liệu được xử lý bằng phương pháp theo sáng chế strongly influences giá trị thuộc tính quan tâm cần dự đoán. Ví dụ, khi thức ăn dùng cho vật nuôi, thức ăn chăn nuôi, hoặc nguyên liệu thức ăn chăn nuôi là nguyên liệu được xử lý bằng phương pháp theo sáng chế, thì giá trị thuộc tính quan tâm cần dự đoán là axit amin, axit béo không no đa chức, và/hoặc khoáng chất dinh dưỡng. Ví dụ, khi phần thức ăn được tiêu hóa và hấp thu, bộ phận cơ thể, hoặc bộ phận cơ thể động vật là nguyên liệu được xử lý bằng phương pháp theo sáng chế, thì giá trị thuộc tính quan tâm cần dự đoán là protein, peptit, và/hoặc axit amin cụ thể. Do đó, phương pháp theo sáng chế không bị giới hạn bất kỳ liên quan đến giá trị thuộc tính quan tâm cần dự đoán. Do đó, giá trị thuộc tính quan tâm có thể là thông số bất kỳ của nguyên liệu, có thể được xác định trực tiếp hoặc gián tiếp bằng phổ hồng ngoại.

Theo sáng chế, mẫu có phổ hồng ngoại được xác định ở bước d) là mẫu có nguồn gốc và/hoặc thành phần chưa biết hoặc có cùng nguồn gốc và/hoặc thành phần như mẫu ở bước a). Tốt hơn nếu, mẫu ở bước d) có nguồn gốc và/hoặc thành phần chưa biết. Theo cách khác, tốt hơn nếu mẫu ở bước d) có cùng nguồn gốc như mẫu ở bước a) nhưng có thành phần chưa biết. Tốt hơn nếu, mẫu có phổ hồng ngoại được xác định ở bước d) là mẫu có nguồn gốc và/hoặc thành phần chưa biết hoặc có cùng nguồn gốc và/hoặc thành phần như quần thể mẫu ở bước a). Tốt hơn nếu, mẫu ở bước d) có nguồn gốc và/hoặc thành phần chưa biết. Theo cách khác, tốt hơn nếu mẫu ở bước d) có cùng nguồn gốc như quần thể mẫu ở bước a) nhưng có thành phần chưa biết.

Do các phổ hồng ngoại có thể được đo theo chế độ truyền qua hoặc phản xạ, mẫu có thể là mẫu chưa được nghiền và/hoặc được nghiền, trong đó mẫu được nghiền có thể có độ kết hạt khác.

Về nguyên tắc, phương pháp theo sáng chế không bị giới hạn ở việc sử dụng giá trị đo khoảng cách. Do đó, giá trị đo khoảng cách thích hợp bất kỳ về nguyên tắc có thể được sử dụng trong phương pháp, ví dụ giá trị đo khoảng cách Euclid, giá trị đo khoảng cách Pearson, hoặc giá trị đo khoảng cách Mahalanobis. Nhìn chung, có thể xác định giá trị đo khoảng cách, được tạo ra dưới dạng hệ số khoảng cách $d(i, j)$ của hai quan sát i và j , từ giá trị đo độ tương tự $s(i, j)$ của hai quan sát i và j , bằng công thức

$$d(i, j) = \sqrt{s(i, j) + s(j, j) - 2s(i, j)}$$

Do đó, cũng có thể thu được giá trị đo khoảng cách một cách gián tiếp, tức là giá trị đo độ tương tự thứ nhất được xác định, sau đó thu nhận giá trị đo khoảng cách từ giá trị đo độ tương tự.

Theo một phương án của phương pháp theo sáng chế giá trị đo khoảng cách là giá trị đo khoảng cách Euclid, giá trị đo khoảng cách Pearson, giá trị đo khoảng cách Mahalanobis, hoặc giá trị đo khoảng cách thu được từ giá trị đo độ tương tự.

Thông thường, phân tích độ tương tự bất kỳ hoạt động với giá trị giới hạn xác định một số loại giới hạn, các quan sát này vẫn có thể chấp nhận được và các quan sát này không còn được chấp nhận. Việc sử dụng giá trị giới hạn nghiêm ngặt cho khoảng cách hoặc giá trị đo độ tương tự thông thường có thể hoạt động tốt trong phòng thí nghiệm và do đó các điều kiện rất dễ lặp lại, với nhân viên được đào tạo tốt và các thiết bị quang phổ, được hiệu chuẩn lý tưởng và là một phần của mạng. Tuy nhiên, việc sử dụng giá trị giới hạn nghiêm ngặt cho khoảng cách hoặc giá trị đo độ tương tự không hoạt động tốt, khi các thiết bị quang phổ độc lập được sử dụng, thường không được vận hành trong điều kiện phòng thí nghiệm. Hơn nữa, việc lựa chọn giá trị giới hạn nghiêm ngặt cho giá trị đo khoảng cách một lần nữa sẽ là vấn đề khá chủ quan, cụ thể khi giá trị giới hạn này chỉ được chọn hoặc sử dụng ngẫu nhiên từ các quy trình khác. Do đó rất có thể rằng việc sử dụng giá trị giới hạn nghiêm ngặt xác định trước cho khoảng cách hoặc giá trị đo độ tương tự sẽ dẫn đến các vấn đề tương tự như các giải pháp đã biết. Tuy nhiên, phương pháp theo sáng chế cũng sẽ thích hợp để dự đoán các giá trị thuộc tính quan tâm bằng cách sử dụng các thiết bị quang phổ độc lập, thường không được vận hành trong điều kiện phòng thí nghiệm.

Các tác giả sáng chế đã phát hiện thấy rằng các vấn đề này được khắc phục bằng cách sử dụng giá trị giới hạn động cho giá trị đo khoảng cách. Cách tiếp cận này của giá trị giới hạn động dẫn đến mức gia tăng từng bước của giá trị giới hạn đo khoảng cách thu được ở bước b4) bằng +1. Do đó, giá trị giới hạn đo khoảng cách “mới” thu được đóng vai trò là cơ sở để đánh giá xem phổ mẫu có phải là giá trị ngoại lệ phổ hay không, theo đó giá trị giới hạn đo khoảng cách thu được ở bước b4) được thay thế bằng giá trị giới hạn đo khoảng cách “mới”. Quy trình này được lặp lại cho đến khi độ chênh lệch giữa hai giá trị đo khoảng cách với các giá trị cao nhất ít nhất bằng 1 và giá trị cao nhất của khoảng cách tối đa bằng 8.

Theo một phương án khác của phương pháp theo sáng chế, bước b) còn bao gồm các bước sau:

b5.1) tăng giá trị giới hạn đo khoảng cách thu được ở bước b4) bằng +1,

b5.2) xác định hai giá trị đo khoảng cách thu được ở bước b5) có các giá trị cao nhất bằng cách sử dụng giá trị giới hạn đo khoảng cách thu được ở bước b5.1),

b5.3) xác định độ chênh lệch giữa các giá trị của giá trị đo khoảng cáchs được xác định ở bước b5.2), và

b5.4) lặp lại các bước b5.1) đến b5.3) với giá trị giới hạn đo khoảng cách thu được ở bước b5.1) cho đến khi độ chênh lệch được xác định ở bước b5.3) ít nhất bằng 1, và giá trị cao nhất của giá trị đo khoảng cách lớn nhất bằng 8.

Mỗi thời điểm các bước b5.1) đến b5.3) được lặp lại, giá trị giới hạn đo khoảng cách từ lần chạy cuối cùng thay thế giá trị giới hạn đo khoảng cách từ lần chạy trước đó, tức là lần chạy trước lần chạy cuối cùng, và được gia tăng +1 trong bước được lặp lại b5.1). Do đó, khi các bước b5.1) đến b5.3) được lặp lại trong lần đầu tiên, giá trị giới hạn đo khoảng cách thu được ở bước b4) được thay thế bằng giá trị giới hạn đo khoảng cách thu được ở bước b5.1) và giá trị giới hạn đo khoảng cách này được gia tăng +1 trong bước được lặp lại b5.1). Khi các bước b5.1) đến b5.3) được lặp lại một lần nữa, giá trị giới hạn đo khoảng cách từ lần chạy cuối cùng thay thế giá trị giới hạn đo khoảng cách từ lần chạy thứ hai đến lần chạy cuối cùng trong bước được lặp lại b5.1) và giá trị giới hạn đo khoảng cách từ lần chạy cuối cùng được gia tăng +1 trong bước được lặp lại b5.1).

Tốt hơn nếu, bước tạo ra hàm dự đoán trên quần thể phổ được làm sạch thu được ở bước b7) bao gồm bước sử dụng tập hợp dữ liệu tham chiếu, thu được từ bước xác định giá trị thuộc tính quan tâm trong mỗi mẫu thu được ở bước a) trong phân tích định lượng.

Theo một phương án khác, phương pháp theo sáng chế còn bao gồm bước a1) xác định các giá trị thuộc tính quan tâm trong mỗi mẫu thu được ở bước a) trong phân tích định lượng để thu được tập hợp dữ liệu tham chiếu.

Thuật ngữ “phân tích định lượng” được sử dụng theo nghĩa rộng nhất đã biết đối với người có hiểu biết trung bình trong lĩnh vực này, tốt hơn nếu trong lĩnh vực hóa phân tích để chỉ phương pháp hóa học và/hoặc vật lý bất kỳ để xác định sự có mặt tuyệt đối hoặc tương đối của một, một số hoặc toàn bộ các thông số cụ thể, ví dụ sự có mặt của các chất

cụ thể, trong mẫu. Ví dụ, khi mẫu truy vấn nằm trong khoảng từ thức ăn dùng cho vật nuôi, thức ăn chăn nuôi hoặc nguyên liệu thức ăn chăn nuôi, thông số cụ thể có thể sự có mặt tuyệt đối hoặc tương đối của một, một số hoặc toàn bộ các axit amin từ danh sách cấu thành từ methionin, cystein, cystin, threonin, leucin, arginin, isoleucin, valin, histidin, phenylalanin, tyrosin, tryptophan, glycin, serin, prolin, alanin, axit aspartic, axit glutamic, tổng lượng của lysin, lượng phản ứng của lysin và tỷ lệ của lượng phản ứng của lysin và tổng lượng của lysin. Thuật ngữ “lượng phản ứng của lysin” được sử dụng trong bản mô tả để chỉ lượng của lysin, thực sự có sẵn cho động vật, cụ thể để tiêu hóa trong động vật. Bằng cách so sánh, thuật ngữ “tổng lượng của lysin” được sử dụng trong bản mô tả để chỉ tổng lượng của lysin, thực sự có sẵn cho động vật, cụ thể để tiêu hóa trong động vật, và lượng của lysin, không sẵn có cho động vật, cụ thể không để tiêu hóa trong động vật. Lượng của lysin không sẵn có cho động vật thường do các phản ứng phân hủy lysin, như phản ứng Maillard đã được đề cập.

Nhiều thức ăn chăn nuôi được chế biến dẫn đến các hư hỏng có thể xuất hiện đối với các axit amin. Điều này có thể tạo ra một số axit amin không có sẵn để sử dụng trong dinh dưỡng. Đây là trường hợp đặc trưng đối với lysin, có nhóm ϵ -amino có thể phản ứng với nhóm carbonyl của các hợp chất khác, ví dụ các đường khử, có mặt trong chế phẩm dinh dưỡng để thu được các hợp chất có thể được hấp thu một phần từ ruột nhưng không có giá trị dinh dưỡng bất kỳ cho động vật. Phản ứng của nhóm ϵ -amino của lysin tự do và/hoặc lysin được gắn kết với protein với các đường khử trong quá trình xử lý nhiệt được gọi là phản ứng Maillard. Phản ứng này tạo ra cả sản phẩm phản ứng Maillard sớm và sản phẩm phản ứng Maillard muộn. Sản phẩm phản ứng Maillard sớm là các dẫn xuất lysin bị biến đổi cấu trúc được gọi là các hợp chất Amadori, trong khi sản phẩm phản ứng Maillard muộn được gọi là melanoidin. Melanoidin không can thiệp vào phân tích thông thường đối với lysin và không ảnh hưởng đến các giá trị tiêu hóa được tính toán. Các sản phẩm này chỉ dẫn đến nồng độ thấp hơn của lysin được hấp thu. Do đó, melanoidin thường không được xác định trong phân tích thông thường của axit amin. Bằng cách so sánh, các hợp chất Amadori can thiệp vào phân tích axit amin và tạo ra nồng độ lysin không chính xác đối với mẫu được phân tích. Lysin được gắn kết trong các hợp chất này được gọi là “lysin phong bế” và không có sẵn về mặt sinh học do lysin này có khả năng chống chịu phản ứng phân hủy bằng enzym đường tiêu hóa bất kỳ.

Hàm lượng lysin phản ứng trong mẫu có thể được xác định bằng cách sử dụng chất phản ứng Sanger, tức là 1-fluoro-2,4-dinitrobenzen (FNDB). Lysin được xác định bằng phương pháp này is do đó cũng được gọi là FDNB-lysin. Chất phản ứng Sanger biến đổi lysin thành dinitrophenyl (DNP)-lysin màu vàng, có thể được chiết xuất và đo bằng phương pháp quang phổ ở bước sóng bằng 435nm hoặc bằng phương pháp sắc ký lỏng hiệu năng cao.

Theo cách khác, hàm lượng lysin phản ứng trong mẫu cũng có thể được xác định bằng phản ứng guanidin hóa bằng cách sử dụng chất phản ứng vừa phải O-metylisourea. Trong phương pháp này O-metylisourea chỉ phản ứng với nhóm ϵ -amino của lysin, nhưng -metylisourea không phản ứng với nhóm α -amino của lysin. Do đó, phản ứng guanidin hóa có thể được sử dụng để xác định lysin tự do và lysin được gắn kết với peptit. Do đó tốt hơn nếu phản ứng guanidin hóa được thực hiện để xác định lysin phản ứng. Phản ứng guanidin hóa của lysin tạo ra homoarginin, tiếp tục được tạo dẫn xuất với ninhydrin và mức độ thay đổi về độ hấp thụ có thể được đo ở bước sóng bằng 570nm. Sau đó, mẫu được tạo dẫn xuất được thủy phân để một lần nữa thu được homoarginin. Bước xác định lysin phản ứng cũng có thể được thực hiện bằng phản ứng guanidin hóa lysin được gắn kết với protein chưa bị hư hỏng trong môi trường kiềm để thu được homoarginin. Phản ứng guanidin hóa kiểu này thường được thực hiện thông qua tác dụng của O-metylisourea (OMIU).

Do phương pháp này dễ sử dụng, nên tốt hơn nếu phản ứng guanidin hóa được sử dụng để bước xác định lysin phản ứng. Phản ứng guanidin hóa bao gồm bước ủ mẫu của nguyên liệu thức ăn chăn nuôi và/hoặc thức ăn chăn nuôi trong O-metylisourea. Tốt hơn nếu, tỷ lệ của O-metylisourea và lysin lớn hơn 1000. Sau đó, mẫu được xử lý thu được từ bước i) được làm khô và phân tích homoarginin, tốt hơn nếu bằng cách sử dụng phương pháp sắc ký lỏng hiệu năng cao trao đổi ion. Sau đó, mẫu này được tạo dẫn xuất với ninhydrin và độ hấp thụ của mẫu được tạo dẫn xuất được đo ở bước sóng bằng 570nm. Sau đó, mẫu này được xử lý bằng phản ứng thủy phân. Khối lượng và khối lượngmol/L của homoarginin trong mẫu được xác định. Cuối cùng, lượng của lysin phản ứng được tính toán từ khối lượngmol/L của homoarginin.

Tuy nhiên, không chỉ lysin bị hư hỏng bởi nhiệt trong quy trình chế biến nguyên liệu thức ăn chăn nuôi và/hoặc thức ăn chăn nuôi mà cả các axit amin khác. Theo phương pháp của sáng chế, các axit amin methionin, cystein, cystin, threonin, leucin, arginin,

isoleucin, valin, histidin, phenylalanin, tyrosin, tryptophan, glycin, serin, prolin, alanin, axit aspartic và axit glutamic được phân tích định lượng trong mẫu của nguyên liệu thức ăn chăn nuôi và/hoặc thức ăn chăn nuôi. Đến mức độ nhất định, các axit amin không chỉ có mặt dưới dạng các hợp chất đơn mà cả dưới dạng oligopeptit, ví dụ dipeptit, tripeptit hoặc peptit bậc cao, được tạo ra trong phản ứng cân bằng từ hai, ba hoặc thậm chí nhiều axit amin. Nhóm amino của axit amin thường là chất ái nhân quá yếu để phản ứng trực tiếp với nhóm carboxyl của một axit amin khác hoặc có mặt ở dạng được proton hóa ($-\text{NH}_3^+$). Do đó, cân bằng của phản ứng này thường ở bên trái trong các điều kiện tiêu chuẩn. Tuy nhiên, phụ thuộc vào các axit amin riêng biệt và điều kiện của dung dịch mẫu chứa một số axit amin được xác định có thể không có mặt dưới dạng các hợp chất đơn nhưng đến mức độ nhất định có mặt dưới dạng oligopeptit, ví dụ dipeptit, tripeptit hoặc peptit bậc cao, được tạo ra từ hai, ba hoặc thậm chí nhiều axit amin. Do đó, mẫu của nguyên liệu thức ăn chăn nuôi và/hoặc thức ăn chăn nuôi cần được xử lý bằng phản ứng thủy phân, tốt hơn nếu phản ứng thủy phân trong môi trường axit hoặc bazơ, bằng cách sử dụng ví dụ axit hydrocloric hoặc bari hydroxit. Để tạo điều kiện cho quá trình phân tách các axit amin tự do và/hoặc định tính và định lượng các axit amin, các axit amin tự do được tạo dẫn xuất bằng chất phản ứng tạo màu, khi cần. Các chất phản ứng tạo màu thích hợp là đã biết đối với người có hiểu biết trung bình trong lĩnh vực này. Sau đó, các axit amin tự do hoặc các axit amin tự do được tạo dẫn xuất được xử lý bằng phân tách sắc ký, trong đó các axit amin khác nhau được phân tách ra khỏi nhau do các thời gian lưu khác nhau do các nhóm chức khác nhau của các axit amin riêng biệt. Các cột sắc ký thích hợp, ví dụ cột trao đổi ion hoặc cột pha đảo, và dung môi rửa giải thích hợp để phân tách bằng sắc ký các axit amin là đã biết đối với người có hiểu biết trung bình trong lĩnh vực này. Cuối cùng, các axit amin được phân tách được xác định trong các dịch rửa giải từ bước sắc ký bằng cách so sánh với chuẩn được hiệu chuẩn, được điều chế để phân tích. Thông thường, các axit amin, được rửa giải từ cột sắc ký, được phát hiện bằng bộ phận phát hiện thích hợp, ví dụ bộ phận phát hiện dẫn điện, bộ phận phát hiện đặc hiệu khối lượng hoặc bộ phận phát hiện huỳnh quang hoặc bộ phận phát hiện UV/VIS phụ thuộc vào các axit amin được tạo dẫn xuất bằng chất phản ứng tạo màu. Bước này tạo ra sắc ký đồ có các diện tích pic và chiều cao pic đối với các axit amin riêng biệt. Bước xác định các axit amin riêng biệt được thực hiện bằng cách so sánh các diện tích pic và chiều cao pic với chuẩn được hiệu chuẩn hoặc đường chuẩn đối với mỗi axit amin. Do cystin ($\text{HO}_2\text{C}(-\text{H}_2\text{N})\text{CH}-\text{CH}_2-\text{S}-\text{S}-\text{CH}_2-\text{CH}(\text{NH}_2)-\text{CO}_2\text{H}$) và

cystein ($\text{HS-CH}_2\text{-CH(NH}_2\text{)-CO}_2\text{H}$) đều được xác định dưới dạng axit cysteic ($\text{HO}_3\text{S-CH}_2\text{-CH(NH}_2\text{)-CO}_2\text{H}$) sau khi thủy phân trong điều kiện axit, nên phân tích định lượng không tạo ra khác biệt bất kỳ giữa hai axit amin do thực tế là cystein rất dễ bị oxy hóa.

Nhìn chung, quy trình được mô tả nêu trên được sử dụng để phân tích định lượng tổng lượng của lysin, cần thiết để xác định tỷ lệ của lượng phản ứng của lysin và tổng lượng của lysin, và phân tích định lượng ít nhất một axit amin được chọn từ nhóm bao gồm methionin, cystein, cystin, threonin, leucin, arginin, isoleucin, valin, histidin, phenylalanin, tyrosin, tryptophan, glycin, serin, prolin, alanin, axit aspartic và axit glutamic.

Điểm mấu chốt nhất trong phân tích định lượng các axit amin là bước điều chế mẫu, khác nhau liên quan đến loại của các thành phần và các axit amin quan tâm chính. Hầu hết axit amin có thể được thủy phân bằng phản ứng thủy phân trong axit hydrocloric (6mol/L) trong thời gian lên đến 24 giờ. Đối với các axit amin chứa lưu huỳnh methionin, cystein, và cystin, phản ứng thủy phân được thực hiện bằng cách oxy hóa, tốt hơn nếu bằng axit performic. Để phân tích định lượng tryptophan, phản ứng thủy phân được thực hiện bằng bari hydroxit (1,5mol/L) trong 20 giờ.

Khi mẫu truy vấn nằm trong khoảng từ thức ăn dùng cho vật nuôi, thức ăn chăn nuôi hoặc nguyên liệu thức ăn chăn nuôi, thông số cụ thể cũng có thể là một, một số hoặc toàn bộ từ danh sách cấu thành từ hoạt tính chất ức chế trypsin, hoạt tính ureaza, độ hòa tan protein trong môi trường kiềm và chỉ số độ phân tán protein của mẫu.

Phân tích định lượng hoạt tính chất ức chế trypsin là trên cơ sở khả năng của các chất ức chế để tạo ra phức chất với enzym trypsin, do đó làm giảm hoạt tính của nó. Trypsin xúc tác phản ứng thủy phân các cơ chất tổng hợp *N*- α -benzoyl-D,L-arginin-*p*-nitroanilit (DL-BAPNA, danh pháp IUPAC *N*-[5-(diaminomethylidenamino)-1-(4-nitroanilino)-1-oxopentan-2-yl]benzylamit) và *N*- α -benzoyl- L-arginin-*p*-nitroanilit (L-BAPNA, danh pháp IUPAC *N*-[5-(diaminomethylidenamino)-1-(4-nitroanilino)-1-oxopentan-2-yl]benzylamit). Phản ứng thủy phân được xúc tác này giải phóng hợp chất có màu vàng *p*-nitroanilin tự do, do đó dẫn đến thay đổi về độ hấp thụ. Hoạt tính trypsin tỷ lệ thuận với cường độ màu vàng. Nồng độ của *p*-nitro anilin có thể được xác định bằng phương pháp quang phổ ở bước sóng bằng 410nm. L-BAPNA thường được sử dụng trong phương pháp ISO 14902 (2001) und DL-BAPNA thường được sử dụng trong phương pháp

AACC 22.40-01 (cải biến của phương pháp ban đầu được phát minh bởi Hamerstrand vào năm 1981).

Trong phương pháp ISO 14902, mẫu trước tiên được nghiền mịn bằng sàng 0,50mm. Trong quá trình nghiền cần tránh phát nhiệt. Mẫu được nghiền được trộn với dung dịch nước kiềm, ví dụ 1g mẫu trong 50mL dung dịch natri hydroxit (0,01 N), và sau đó dung dịch, hỗn dịch, thể phân tán hoặc nhũ tương thu được được bảo quản trong thời gian lên đến khoảng 24 giờ ở nhiệt độ tối đa bằng 4°C. Hỗn hợp thu được có độ pH nằm trong khoảng từ 9 đến 10, đặc biệt nằm trong khoảng từ 9,4 đến 9,6. Dung dịch thu được được pha loãng bằng nước và để yên. Mẫu của dung dịch này, ví dụ 2mL, được lấy và pha loãng như được thể hiện bằng hoạt tính chất ức chế trypsin giả định hoặc gần đúng trước đó của nó sao cho 1mL dung dịch pha loãng sẽ ức chế từ 40% đến 60% phản ứng enzym. Dung dịch làm việc trypsin, ví dụ 1mL, được bổ sung vào hỗn hợp của L-BAPNA, nước và dung dịch chiết xuất mẫu pha loãng thích hợp, ví dụ 5mL L-BAPNA, 2mL nước cất và 1mL dung dịch chiết xuất mẫu pha loãng thích hợp. Sau đó các mẫu được ủ chính xác trong 10 phút ở nhiệt độ 37°C. Phản ứng được dừng bằng cách bổ sung 1mL axit axetic (30%). Mẫu trắng được điều chế như nêu trên, nhưng trypsin được bổ sung sau axit axetic. Sau khi ly tâm ở tốc độ bằng 2,5g, độ hấp thụ được đo ở bước sóng bằng 410nm.

Trong phương pháp AACC 22-40,01, mẫu trước tiên được nghiền mịn bằng sàng 0,15mm. Trong quá trình nghiền cần tránh phát nhiệt. Mẫu được nghiền được trộn với dung dịch nước kiềm, ví dụ 1g mẫu trong 50mL dung dịch natri hydroxit (0,01 N), và khuấy từ từ trong 3 giờ ở nhiệt độ 20°C. Độ pH của dung dịch, hỗn dịch, thể phân tán hoặc nhũ tương thu được cần nằm trong khoảng từ 8 đến 11, tốt hơn nếu nằm trong khoảng từ 8,4 đến 10. Dung dịch, hỗn dịch, thể phân tán hoặc nhũ tương thu được được pha loãng bằng nước, lắc và để yên. Mẫu của dung dịch này, ví dụ 2mL, được lấy và pha loãng như được thể hiện bằng hoạt tính chất ức chế trypsin giả định hoặc gần đúng trước đó của nó sao cho 1mL dung dịch pha loãng sẽ ức chế từ 40% đến 60% phản ứng enzym. Dung dịch làm việc trypsin, ví dụ 2 ml, được bổ sung vào hỗn hợp của D,L-BAPNA, nước và dung dịch chiết xuất mẫu pha loãng thích hợp, ví dụ 5mL D,L-BAPNA, 1mL nước cất và 1mL dung dịch chiết xuất mẫu pha loãng thích hợp. Sau đó các mẫu được ủ chính xác trong 10 phút ở nhiệt độ 37°C. Phản ứng được dừng bằng cách bổ sung 1mL axit axetic (30%). Mẫu trắng được

điều chế như nêu trên, nhưng trypsin được bổ sung sau axit axetic. Sau khi ly tâm ở tốc độ bằng 2,5g, độ hấp thụ được đo ở bước sóng bằng 410nm.

Không phụ thuộc vào phương pháp được sử dụng, hoạt tính chất ức chế trypsin được tính toán theo đơn vị mg chất ức chế trypsin/1 g trypsin, bằng công thức sau:

$$i = \frac{(Ar - Abr) - (As - Abs)}{(Ar - Abr)}$$

i = tỷ lệ phần trăm ức chế (%);

Ar = độ hấp thụ của dung dịch chứa chất chuẩn;

Abr = độ hấp thụ của mẫu trắng chứa chất chuẩn;

As = độ hấp thụ của dung dịch chứa mẫu;

Abs = độ hấp thụ của mẫu trắng chứa mẫu;

$$TIA = \frac{i}{100\%} \times \frac{m1 \times f1 \times f2}{m0}$$

TIA = hoạt tính chất ức chế trypsin (mg/g);

i = tỷ lệ phần trăm ức chế (%);

$m0$ = khối lượng của mẫu thử nghiệm (g);

$m1$ = khối lượng của trypsin (g);

$f1$ = hệ số pha loãng của chiết xuất mẫu; và

$f2$ = hệ số biến đổi trên cơ sở độ tinh khiết của trypsin.

Một đơn vị trypsin được định nghĩa là lượng của enzym, sẽ làm tăng 0,01 đơn vị độ hấp thụ ở bước sóng bằng 410nm sau 10 phút phản ứng đối với mỗi 1mL thể tích phản ứng. Hoạt tính chất ức chế trypsin được định nghĩa là số lượng của đơn vị trypsin được ức chế (TUI). TUI tính trên 1mL được tính toán bằng cách sử dụng công thức

$$TUI [ml] = \frac{A_{blank} - A_{sample}}{0.01 \times V_{dl.smp.}}$$

trong đó:

$A_{mẫu trắng}$ = độ hấp thụ mẫu trắng

$A_{mẫu}$ = độ hấp thụ mẫu

$V_{dl. smp.}$ = thể tích của dung dịch mẫu được pha loãng (mL).

TUI thu được được vẽ đồ thị theo các thể tích của dung dịch mẫu được pha loãng, trong đó giá trị ngoại suy của thể tích chất ức chế đến 0mL tạo ra TUI cuối cùng [mL]. Cuối cùng, TUI tính trên 1g mẫu được tính toán bằng công thức

$$TUI [g] = TUI [ml^{-1}] \times d \times 50$$

trong đó d = hệ số pha loãng (thể tích cuối cùng chia cho lượng được lấy).

Các kết quả của phương pháp phân tích này cần không vượt quá 10% giá trị trung bình đối với các mẫu lặp lại.

Enzym ureaza xúc tác phản ứng phân hủy ure thành amoniac và cacbon dioxit. Do ureaza có tự nhiên trong đậu tương, nên phân tích định lượng enzym này là thử nghiệm phổ biến nhất để đánh giá chất lượng của đậu tương chế biến. Tốt hơn nếu, phân tích định lượng ureaza được thực hiện theo phương pháp của ISO 5506 (1988) hoặc AOCS Ba 9-58. Phương pháp của AOCS Ba 9-58 xác định hoạt tính dư của ureaza dưới dạng chỉ báo gián tiếp để đánh giá xem các chất ức chế proteaza đã được phân hủy trong quy trình chế biến nguyên liệu thức ăn chăn nuôi và/hoặc thức ăn chăn nuôi hay chưa. Hoạt tính dư của ureaza được đo dưới dạng mức độ gia tăng độ pH trong thử nghiệm là kết quả của quá trình giải phóng hợp chất kiềm ammonia vào môi trường. Mức độ gia tăng được khuyến cáo của độ pH nằm trong khoảng từ 0,01 đến 0,35 đơn vị (NOPA, 1997). Thông thường, phân tích định lượng hoạt tính ureaza của nguyên liệu thức ăn chăn nuôi và/hoặc thức ăn chăn nuôi được thực hiện tương tự: Trước tiên, dung dịch chứa urea trong dung dịch đệm bao gồm NaH_2PO_4 và KH_2PO_4 được điều chế, ví dụ 30g ure được bổ sung vào 1L dung dịch đệm cấu thành từ 4,45g Na_2HPO_4 và 3,4g KH_2PO_4 và độ pH của hỗn hợp thu được được đo. Sau đó, mẫu của nguyên liệu thức ăn chăn nuôi và/hoặc thức ăn chăn nuôi, ví dụ 0,2g mẫu đậu tương, được bổ sung vào dung dịch này. Ống thử nghiệm hoặc cốc thử nghiệm chứa dung dịch, hỗn dịch, thể phân tán, hoặc nhũ tương thu được được đặt vào bể nước, ví dụ ở nhiệt độ $30 \pm 5^\circ C$, tốt hơn nếu $30^\circ C$, trong 20 đến 40 phút, tốt hơn nếu 30 phút. Cuối cùng, độ pH của dung dịch, hỗn dịch, thể phân tán, hoặc nhũ tương được đặt vào được đo, so sánh với độ pH của dung dịch ure ban đầu, và độ chênh lệch được tạo ra dưới dạng mức độ gia tăng pH.

Độ hòa tan của các protein trong môi trường kiềm, sau đây cũng được gọi là độ hòa tan của các protein trong dung dịch kiềm hoặc độ hòa tan trong môi trường kiềm của các protein, là phương pháp hữu hiệu để phân biệt các sản phẩm được chế biến quá mức với các sản phẩm được chế biến chính xác, ví dụ theo DIN EN ISO 14244.

Độ hòa tan của các protein trong môi trường kiềm hoặc độ hòa tan trong môi trường kiềm của các protein bao gồm bước xác định tỷ lệ phần trăm của protein được hòa tan trong dung dịch kiềm. Trước khi hòa tan mẫu có khối lượng đã biết của nguyên liệu thức ăn chăn nuôi và/hoặc thức ăn chăn nuôi, hàm lượng nitơ của mẫu có khối lượng xác định được xác định bằng cách sử dụng phương pháp chuẩn để xác định nitơ, như phương pháp Kjeldahl hoặc phương pháp Dumas. Hàm lượng nitơ xác định được gọi là hàm lượng nitơ tổng số. Sau đó, mẫu có cùng khối lượng và nguồn gốc được tạo hỗn dịch trong dung dịch kiềm có nồng độ xác định, tốt hơn nếu trong dung dịch kiềm hydroxit, cụ thể trong dung dịch kali hydroxit. Mẫu có thể tích xác định của hỗn dịch thu được được lấy và ly tâm. Mẫu có thể tích xác định của hỗn dịch thu được được lấy một lần nữa. Hàm lượng nitơ trong phân đoạn lỏng này được xác định bằng cách sử dụng phương pháp chuẩn để xác định nitơ, như phương pháp Kjeldahl hoặc phương pháp Dumas. Hàm lượng nitơ xác định được so sánh với hàm lượng nitơ tổng số và biểu diễn dưới dạng tỷ lệ phần trăm của hàm lượng nitơ ban đầu của mẫu.

Dung dịch kiềm thông thường để xác định độ hòa tan trong môi trường kiềm của các protein có độ pH bằng 12,5, ví dụ dung dịch chứa kali hydroxit có nồng độ bằng 0,036mol/L hoặc 0,2% khối lượng. Ở bước ii) 1,5g mẫu đậu tương có thể được bổ sung vào 75mL dung dịch kali hydroxit, sau đó khuấy ở tốc độ bằng 8500 vòng/phút trong 20 phút ở nhiệt độ 20°C. Sau đó, mẫu có thể tích xác định, ví dụ khoảng 50mL hỗn dịch, dung dịch, thể phân tán hoặc nhũ tương thu được được lấy và ly tâm ngay lập tức ở tốc độ bằng 2500g trong 15 phút. Sau đó, mẫu có thể tích xác định, ví dụ 10mL dịch nổi của hỗn dịch, dung dịch, thể phân tán hoặc nhũ tương thu được được lấy và hàm lượng của nitơ trong mẫu có thể tích xác định này được xác định bằng các phương pháp chuẩn để xác định nitơ, như phương pháp Kjeldahl hoặc phương pháp Dumas. Cuối cùng, các kết quả được biểu diễn dưới dạng tỷ lệ phần trăm của hàm lượng nitơ của mẫu.

Bước xác định chỉ số độ phân tán protein (PDI) đo độ hòa tan của các protein trong nước sau khi trộn mẫu với nước. Phương pháp này cũng bao gồm bước xác định hàm lượng

nitơ trong mẫu có khối lượng đã biết, thường được thực hiện theo quy trình tương tự như trong phân tích hóa học ướt của các protein. Hàm lượng nitơ thu được cũng được gọi là hàm lượng nitơ tổng số. Hơn nữa, phương pháp cũng bao gồm bước điều chế hỗn dịch của mẫu có cùng khối lượng như trong bước xác định hàm lượng nitơ được tạo hỗn dịch trong nước, thường được thực hiện bằng cách sử dụng thiết bị trộn tốc độ cao. Hỗn dịch thu được được lọc và dịch lọc được xử lý bằng cách ly tâm. Hàm lượng nitơ trong thu được dịch nổi được xác định bằng cách sử dụng một lần nữa phương pháp chuẩn để xác định, như phương pháp Kjeldahl hoặc phương pháp Dumas, được mô tả nêu trên. Hàm lượng nitơ thu được cũng được gọi là hàm lượng nitơ trong dung dịch. Chỉ số độ phân tán protein cuối cùng được tính toán dưới dạng tỷ lệ của hàm lượng nitơ trong dung dịch và hàm lượng nitơ tổng số và biểu diễn dưới dạng tỷ lệ phần trăm của hàm lượng nitơ ban đầu của mẫu.

Do các giá trị để chỉ số độ phân tán protein tăng với kích cỡ hạt giảm, các kết quả thu được ở bước xác định chỉ số độ phân tán protein phụ thuộc vào kích cỡ hạt của mẫu. Do đó tốt hơn nếu nghiền mẫu được xử lý bằng bước xác định chỉ số độ phân tán protein, cụ thể bằng kích cỡ rây 1mm.

Quy trình được mô tả nêu trên là theo phương pháp chính thức Ba 10-65 của American Oil Chemists' Society (A.O.C.S.), theo đó tốt hơn nếu bước xác định chỉ số độ phân tán protein được thực hiện. Hàm lượng nitơ của ví dụ mẫu đậu tương được xác định bằng các phương pháp chuẩn để xác định nitơ, như phương pháp Kjeldahl hoặc phương pháp Dumas. Mẫu có thể tích xác định, ví dụ 20g mẫu đậu tương được đặt vào thiết bị trộn, và nước khử ion, ví dụ 300mL, được bổ sung vào ở nhiệt độ 25°C, sau đó khuấy, ví dụ ở tốc độ bằng 8500 vòng/phút trong 10 phút. Hỗn dịch, dung dịch, thể phân tán hoặc nhũ tương thu được được lọc và dung dịch, thể phân tán hoặc nhũ tương thu được được ly tâm, ví dụ ở tốc độ bằng 1000g trong 10 phút. Cuối cùng, hàm lượng nitơ trong dịch nổi được xác định bằng các phương pháp chuẩn để xác định nitơ, như phương pháp Kjeldahl hoặc phương pháp Dumas.

Tập hợp dữ liệu tham chiếu thu được từ bước a1) được sử dụng để xác định mối tương quan dữ liệu với quần thể phổ được làm sạch thu được ở bước b1) để tạo ra hàm dự đoán ở bước c). Trong trường hợp của sự có mặt tuyệt đối hoặc tương đối của một, một số hoặc toàn bộ các chất truy vấn trong mẫu, mối tương quan dữ liệu liên quan đến nồng độ của các chất truy vấn trong mẫu.

Theo phương án ưu tiên của phương pháp theo sáng chế, bước tạo ra hàm dự đoán ở bước c) bao gồm bước phân tích tập hợp dữ liệu tham chiếu thu được ở bước a1) và quần thể phổ được làm sạch thu được ở bước b7) đối với mỗi tương quan dữ liệu để thu được hàm dự đoán.

Chất lượng dự đoán đạt được bằng phương pháp theo sáng chế có thể được cải thiện bằng cách xử lý hàm dự đoán được tạo ra ở bước c) bằng bước kiểm chứng.

Theo một phương án khác của phương pháp theo sáng chế, bước tạo ra hàm dự đoán ở bước c) bao gồm bước kiểm chứng hàm dự đoán này.

Về nguyên tắc, phương pháp theo sáng chế không bị giới hạn bất kỳ liên quan đến loại của quy trình kiểm chứng. Do đó, quy trình kiểm chứng thích hợp bất kỳ có thể được sử dụng trong phương pháp theo sáng chế. Tuy nhiên, theo sáng chế, tốt hơn nếu sử dụng bước kiểm chứng chéo chia đôi, bước kiểm chứng chéo k lần và/hoặc bước kiểm chứng dựa trên dữ liệu kiểm định vòng. Bước kiểm chứng dựa trên dữ liệu kiểm định vòng là bước kiểm chứng dựa trên các phổ bên ngoài và dữ liệu tham chiếu bên ngoài tương ứng. Tùy chọn này có ưu điểm là quần thể hiệu chuẩn không chỉ được đánh giá liên quan đến hiệu suất của nó trên thiết bị quang phổ chính, tức là thiết bị quang phổ được sử dụng để tạo ra hàm dự đoán. Hơn nữa, tùy chọn này cũng cho phép đánh giá hiệu suất của quần thể hiệu chuẩn liên quan đến mạng với nhiều thiết bị quang phổ khác, ví dụ hàng trăm thiết bị, cũng sử dụng hàm hiệu chuẩn. Do đó tốt hơn nếu sử dụng bước kiểm chứng dựa trên dữ liệu kiểm định vòng.

Theo phương án ưu tiên của phương pháp theo sáng chế, bước kiểm chứng là bước kiểm chứng chéo chia đôi, bước kiểm chứng chéo k lần và/hoặc bước kiểm chứng dựa trên dữ liệu kiểm định vòng.

Theo sáng chế, các thuật ngữ bước kiểm chứng chéo chia đôi, bước kiểm chứng chéo k lần và bước kiểm chứng dựa trên dữ liệu kiểm định vòng được sử dụng như đã biết đối với người có hiểu biết trung bình trong lĩnh vực này.

Ví dụ, khi bước kiểm chứng là bước kiểm chứng chéo chia đôi, bước c) của phương pháp theo sáng chế còn bao gồm các bước sau:

cV1) phân chia dữ liệu tham chiếu thu được ở bước a1) và quần thể phổ được làm sạch thu được ở bước b7) thành tập huấn luyện S0 và tập thử nghiệm S1, trong đó kích cỡ của tập thử nghiệm S1 nhỏ hơn kích cỡ của tập huấn luyện S0,

cV2) tạo ra hàm dự đoán sơ bộ trên tập huấn luyện S0 thu được ở bước cV1),

cV3) dự đoán giá trị thuộc tính quan tâm bằng cách gán hàm dự đoán sơ bộ thu được ở bước cV2) trên quần thể phổ được làm sạch,

cV4) tính toán sai số giá trị trung bình dự đoán của hàm dự đoán sơ bộ thu được ở bước cV2) với các giá trị thuộc tính dự đoán thu được ở bước cV3) và dữ liệu tham chiếu tương ứng thu được ở bước cV1),

cV5) lặp lại các bước cV1) đến cV5), khi sai số giá trị trung bình dự đoán của hàm dự đoán sơ bộ nằm ngoài giới hạn, hoặc thực hiện bước cV6), khi sai số giá trị trung bình dự đoán của hàm dự đoán sơ bộ nằm trong giới hạn, và

cV6) chấp nhận hàm dự đoán sơ bộ thu được ở bước cV2) làm hàm dự đoán.

Ví dụ, khi bước kiểm chứng là bước kiểm chứng chéo k lần, bước c) của phương pháp theo sáng chế còn bao gồm các bước sau:

cV1) phân chia tập quần thể phổ được làm sạch thu được ở bước b7) và dữ liệu tham chiếu tương ứng thu được ở bước a1) thành n tập con có kích cỡ bằng nhau với n là số nguyên với giá trị cực tiểu ít nhất bằng 2 và giá trị tối đa nhỏ hơn số lượng của các phổ, trong đó một trong số các tập con được lấy làm tập thử nghiệm và các tập con còn lại được lấy làm các tập huấn luyện,

cV2) tạo ra hàm dự đoán sơ bộ trên tập huấn luyện thu được ở bước cV1),

cV3) dự đoán giá trị thuộc tính quan tâm bằng cách gán hàm dự đoán sơ bộ thu được ở bước cV2) trên quần thể phổ được làm sạch,

cV4) tính toán sai số giá trị trung bình dự đoán của hàm dự đoán sơ bộ với các giá trị thuộc tính dự đoán thu được ở bước cV3) và dữ liệu tham chiếu tương ứng thu được ở bước cV1),

cV5) lặp lại các bước cV1) đến cV5), khi sai số giá trị trung bình dự đoán của hàm dự đoán sơ bộ nằm ngoài giới hạn, hoặc thực hiện bước cV6), khi sai số giá trị trung bình dự đoán của hàm dự đoán sơ bộ nằm trong giới hạn,

cV6) chấp nhận hàm dự đoán sơ bộ thu được ở bước cV2) làm hàm dự đoán và thực hiện các bước cV2) đến cV6) k lần với mỗi tập trong số n tập con được sử dụng một lần làm tập thử nghiệm để thu được k hàm dự đoán, và

cV7) tính giá trị trung bình các thông số của mỗi hàm dự đoán được chấp nhận thu được ở bước cV6) để thu được hàm hiệu chuẩn.

Ví dụ, khi bước kiểm chứng là bước kiểm chứng dựa trên dữ liệu kiểm định vòng, bước c) của phương pháp theo sáng chế còn bao gồm các bước sau:

cV1) xác định tập phổ tham chiếu bên ngoài và dữ liệu tham chiếu bên ngoài tương ứng,

cV2) tạo ra hàm dự đoán sơ bộ trên quần thể phổ được làm sạch thu được ở bước b7) và dữ liệu tham chiếu thu được ở bước a1),

cV3) dự đoán giá trị thuộc tính quan tâm bằng cách gán hàm dự đoán sơ bộ thu được ở bước cV2) on các phổ tham chiếu bên ngoài thu được ở bước cV1),

cV4) tính toán sai số giá trị trung bình dự đoán của hàm dự đoán sơ bộ thu được ở bước cV2) với các giá trị thuộc tính dự đoán thu được ở bước cV3) và dữ liệu tham chiếu bên ngoài tương ứng thu được ở bước cV1),

cV5) lặp lại các bước cV1) đến cV5), khi sai số giá trị trung bình dự đoán của hàm dự đoán sơ bộ nằm ngoài giới hạn, hoặc thực hiện bước cV6), khi sai số giá trị trung bình dự đoán của hàm dự đoán sơ bộ nằm trong giới hạn, và

cV6) chấp nhận hàm dự đoán sơ bộ thu được ở bước cV2) làm hàm dự đoán.

Phụ thuộc vào loại quy trình kiểm chứng, sai số giá trị trung bình dự đoán cũng được gọi là sai số trung bình căn bậc hai bình phương của kiểm chứng chéo (RMSECV), khi bước kiểm chứng là bước kiểm chứng bên trong, cụ thể bước kiểm chứng chéo, và được tính toán bằng công thức

$$RMSECV = \sqrt{\frac{1}{M} \sum_{i=1}^M (y_i^{measured} - y_i^{predicted})^2}$$

trong đó M = số lượng của các giá trị thuộc tính dự đoán,

$y_i^{measured}$ = giá trị thuộc tính đo được,

$$y_i^{predicted} = \text{giá trị thuộc tính dự đoán.}$$

Nhìn chung, bước kiểm chứng hàm dự đoán thu được bằng phương pháp theo sáng chế hoặc phương án bất kỳ của nó đã tạo ra hàm dự đoán hiệu dụng và tin cậy. Tuy nhiên, trong một số ít trường hợp, bước kiểm chứng này có thể thích hợp để tiếp tục đánh giá hàm dự đoán được tạo ra trong phương pháp theo sáng chế hoặc phương án bất kỳ của nó. Một cách tiếp cận hữu ích là vẽ đồ thị, các giá trị thuộc tính được dự đoán đối với quần thể phổ được làm sạch thu được ở bước b7), và hàm dự đoán thu được ở bước c) dưới dạng đường hồi quy thành đồ thị giá trị dự đoán so với giá trị đo được để thu được đồ thị phân tán. Tuy nhiên, các giá trị thuộc tính của dữ liệu tham chiếu thu được ở bước a1) cũng được vẽ đồ thị thành đồ thị này. Sau đó, đường phân giác được vẽ đồ thị thành thu được đồ thị phân tán để chia mặt phẳng của đồ thị này thành hai mặt phẳng hình tam giác có kích cỡ bằng nhau. Về mặt lý thuyết, hàm dự đoán và đường phân giác được vẽ đồ thị sẽ tương đẳng, khi hàm dự đoán tối ưu. Tuy nhiên, trong thực tế, điều này hiếm khi xảy ra. Tiếp theo, các khoảng cách của mỗi giá trị thuộc tính dự đoán đến đường phân giác trong đồ thị phân tán được xác định, và hạng tổng thể của các khoảng cách này được tạo ra, với các khoảng cách cao nhất ở đỉnh của hạng. Các phổ tương ứng với ít nhất ba khoảng cách được đánh giá ở đỉnh, tức là các giá trị thuộc tính dự đoán, được loại bỏ ra khỏi quần thể phổ được làm sạch thu được ở bước b7). Sau đó, hàm dự đoán thu được ở bước c) được hiệu chỉnh để bù giá trị loại bỏ của các phổ. Cách tiếp cận này cho phép tiếp tục tinh chỉnh hàm dự đoán. Kết quả của bước tinh chỉnh này có thể dễ dàng được kiểm soát khi cách tiếp cận này được lặp lại từng bước: mỗi lần lặp lại di chuyển hàm dự đoán gần hơn với đường phân giác.

Theo một phương án ưu tiên khác của phương pháp theo sáng chế, bước c) còn bao gồm các bước sau:

c1) vẽ đồ thị các giá trị thuộc tính được dự đoán đối với quần thể phổ được làm sạch thu được ở bước b7), và hàm dự đoán thu được ở bước c) dưới dạng đường hồi quy thành đồ thị giá trị dự đoán so với giá trị đo được để thu được đồ thị phân tán,

c2) vẽ đồ thị đường phân giác giữa các trục của đồ thị thu được ở bước c1),

c3) xác định các khoảng cách của mỗi giá trị thuộc tính dự đoán thu được ở bước c1) đến đường phân giác trong đồ thị phân tán thu được ở bước c2),

c4) tạo ra hạng tổng thể của các khoảng cách thu được ở bước c3), với các khoảng cách cao nhất ở đỉnh của hạng, và loại bỏ các phổ tương ứng với ít nhất ba khoảng cách được đánh giá ở đỉnh ra khỏi quần thể phổ được làm sạch thu được ở bước b7), và

c5) hiệu chỉnh hàm dự đoán thu được ở bước c) để bù giá trị loại bỏ của các phổ ở bước c4).

Thuật ngữ “hiệu chỉnh hàm dự đoán được sử dụng trong bản mô tả để chỉ bước tạo ra hàm dự đoán mới trên quần thể phổ được làm sạch thu được từ bước loại bỏ bổ sung các phổ ở bước c4) nêu trên.

Trong một số trường hợp hai hoặc nhiều phổ tương ứng với các khoảng cách được xếp hạng cao nhất, cần được loại bỏ ra khỏi quần thể phổ được làm sạch thu được ở bước b7), có thể ở vùng lân cận (chặt chẽ) trong đồ thị phân tán. Tuy nhiên, giá trị loại bỏ của hai hoặc nhiều phổ ở vùng lân cận (chặt chẽ) có thể ảnh hưởng quá mức đến độ dốc của hàm dự đoán được hiệu chỉnh. Trong trường hợp xấu nhất, độ dốc của hàm dự đoán bị ảnh hưởng đáng kể, ví dụ hiệu chỉnh hàm dự đoán dẫn đến độ dốc cao hơn hoặc thấp hơn trước.

Hiệu ứng này có thể được ngăn ngừa hoặc ít nhất được giảm đáng kể bằng cách gán trọng số cho hai hoặc nhiều phổ tương ứng với các khoảng cách được xếp hạng cao nhất, cần được loại bỏ ra khỏi quần thể phổ được làm sạch thu được ở bước b7), khi chúng ở vùng lân cận (chặt chẽ) trong đồ thị phân tán. Vì mục đích này, trước tiên, đường thẳng đứng được vẽ đồ thị qua đường phân giác ở một nửa chiều dài của đường phân giác này, sao cho mỗi mặt phẳng ở phía trên và phía dưới đường phân giác được chia thành hai mặt phẳng, trong đó các mặt phẳng tương ứng được phản chiếu ở đường phân giác có kích cỡ bằng nhau. Nhìn chung, bốn mặt phẳng được tạo ra bằng cách chia mặt phẳng giữa các trục của đồ thị phân tán bằng đường phân giác góc và đường thẳng đứng được vẽ đồ thị qua đường phân giác ở một nửa chiều dài của nó: hai hình tam giác có kích cỡ bằng nhau và hai hình vuông có kích cỡ bằng nhau, với mỗi một hình tam giác và một hình vuông ở phía trên và phía dưới đường phân giác. Tiếp theo, hai hoặc nhiều phổ tương ứng với các khoảng cách được xếp hạng cao nhất, cần được loại bỏ ra khỏi quần thể phổ được làm sạch thu được ở bước b7), được gán trọng số bằng hệ số bằng 0,5, khi chúng có mặt trong cùng mặt phẳng được tạo ra bởi đường phân giác và đường thẳng đứng.

Do đó, theo một phương án ưu tiên khác, quy trình theo sáng chế còn bao gồm các dấu hiệu sau:

-ở bước c2) đường thẳng đứng được vẽ đồ thị qua đường phân giác ở một nửa chiều dài của đường phân giác này, để chia mặt phẳng giữa các trục của đồ thị phân tán thành bốn mặt phẳng, và

-ở bước c4) các phổ tương ứng với ít nhất hai khoảng cách được xếp hạng cao nhất được gán trọng số bằng hệ số bằng 0,5, khi chúng có mặt trong cùng mặt phẳng.

Về nguyên tắc, cách tiếp cận được mô tả nêu trên có thể được thực hiện cho đến khi hàm dự đoán và đường phân giác tương đẳng. Trong thực tế, quy trình thu được ở các bước c1) đến c5) sẽ được lặp lại cho đến khi độ hội tụ của hàm dự đoán với đường phân giác không được cải thiện hơn nữa, hoặc nói cách khác cho đến khi đồ thị của hàm dự đoán được hiệu chỉnh thu được dưới dạng đường hồi quy càng gần càng tốt với đường phân giác.

Theo một phương án ưu tiên khác, quy trình theo sáng chế còn bao gồm bước c6) lặp lại các bước c1) đến c5) với hàm dự đoán được hiệu chỉnh và quần thể phổ được làm sạch thu được từ bước hiệu chỉnh trước đó cho đến khi đồ thị của hàm dự đoán được hiệu chỉnh thu được dưới dạng đường hồi quy càng gần càng tốt với đường phân giác.

Ngoài ra, quy trình thu được ở các bước c1) đến c5) có hoặc không có bước c6), cũng có thể tiếp tục cải thiện chất lượng của hàm dự đoán thu được ở bước c) bằng khoảng tin cậy. Trước tiên, quy trình này bao gồm bước vẽ đồ thị các giá trị thuộc tính được dự đoán được vẽ đồ thị đối với quần thể phổ được làm sạch thu được ở bước b7), và hàm dự đoán thu được ở bước c) dưới dạng đường hồi quy thành đồ thị giá trị dự đoán so với giá trị đo được để thu được đồ thị phân tán. Tiếp theo, tương tự trong quy trình được mô tả nêu trên, đường phân giác được vẽ đồ thị giữa hai trục của đồ thị phân tán, và khoảng tin cậy có chiều rộng xác định trước được vẽ đồ thị xung quanh đường phân giác. Sau đó, các phổ tương ứng với các điểm nằm ngoài khoảng tin cậy được loại bỏ ra khỏi quần thể phổ được làm sạch thu được ở bước b7) và hàm dự đoán thu được ở bước c) được xử lý bằng bước hiệu chỉnh để bù giá trị loại bỏ của các phổ. Theo sáng chế, khoảng tin cậy có chiều rộng nằm trong khoảng từ $+5 \times \text{RMSECV}$ đến $-5 \times \text{RMSECV}$, tốt hơn nếu nằm trong khoảng từ $+4 \times \text{RMSECV}$ đến $-4 \times \text{RMSECV}$, nằm trong khoảng từ $+3 \times \text{RMSECV}$ đến $-3 \times \text{RMSECV}$, nằm trong khoảng từ $+2 \times \text{RMSECV}$ đến $-2 \times \text{RMSECV}$, nằm trong khoảng từ $+1 \times \text{RMSECV}$ đến $-1 \times \text{RMSECV}$ hoặc chiều rộng nằm trong khoảng từ số thực dương bất kỳ đến số thực âm tương ứng nằm trong khoảng từ $+5 \times \text{RMSECV}$ đến $-5 \times \text{RMSECV}$, xung quanh giá trị tương ứng của đường phân giác góc, trong đó RMSECV là sai số trung bình căn bậc hai bình phương của kiểm chứng chéo như được định nghĩa nêu trên.

Theo một phương án khác phương pháp theo sáng chế còn bao gồm các bước sau:

c7) vẽ đồ thị các giá trị thuộc tính được dự đoán đối với quần thể phổ được làm sạch thu được ở bước b7), và hàm dự đoán thu được ở bước c) dưới dạng đường hồi quy thành đồ thị giá trị dự đoán so với giá trị đo được để thu được đồ thị phân tán,

c8) vẽ đồ thị đường phân giác giữa hai trục của đồ thị thu được ở bước c7),

c9) vẽ đồ thị khoảng tin cậy với chiều rộng xác định trước xung quanh đường phân giác thành đồ thị thu được ở bước c8),

c10) loại bỏ các phổ tương ứng với các điểm nằm ngoài khoảng tin cậy trong đồ thị thu được ở bước c9) ra khỏi quần thể phổ được làm sạch thu được ở bước b7), và

c11) hiệu chỉnh hàm dự đoán thu được ở bước c) để bù giá trị loại bỏ của các phổ ở bước c10).

Thuật ngữ “hiệu chỉnh hàm dự đoán được sử dụng trong bản mô tả để chỉ bước tạo ra hàm dự đoán mới trên quần thể phổ được làm sạch thu được từ bước loại bỏ bỏ sung các phổ ở bước c10) nêu trên.

Một lần nữa, trong một số trường hợp hai hoặc nhiều phổ tương ứng với các khoảng cách được xếp hạng cao nhất, cần được loại bỏ ra khỏi quần thể phổ được làm sạch thu được ở bước b7), có thể ở vùng lân cận (chặt chẽ) trong đồ thị phân tán. Tuy nhiên, giá trị loại bỏ của hai hoặc nhiều phổ ở vùng lân cận (chặt chẽ) có thể ảnh hưởng quá mức đến độ dốc của hàm dự đoán được hiệu chỉnh. Trong trường hợp xấu nhất độ dốc của hàm dự đoán bị ảnh hưởng đáng kể, ví dụ hiệu chỉnh hàm dự đoán dẫn đến độ dốc cao hơn hoặc thấp hơn trước.

Hiệu ứng này có thể được ngăn ngừa hoặc ít nhất được giảm đáng kể bằng cách gán trọng số cho hai hoặc nhiều phổ tương ứng với các khoảng cách được xếp hạng cao nhất, cần được loại bỏ ra khỏi quần thể phổ được làm sạch thu được ở bước b7), khi chúng ở vùng lân cận (chặt chẽ) trong đồ thị phân tán. Vì mục đích này, trước tiên, đường thẳng đứng được vẽ đồ thị qua đường phân giác ở một nửa chiều dài của đường phân giác này, sao cho các mặt phẳng này ở phía trên và phía dưới đường phân giác được chia thành hai mặt phẳng mỗi, trong đó các mặt phẳng tương ứng được phản chiếu ở đường phân giác có kích cỡ bằng nhau. Nhìn chung, bốn mặt phẳng được tạo ra bằng cách chia mặt phẳng giữa các trục của đồ thị phân tán với đường phân giác góc và đường thẳng đứng được vẽ đồ

thị qua đường phân giác ở một nửa chiều dài của nó: hai hình tam giác có kích cỡ bằng nhau và hai hình vuông có kích cỡ bằng nhau, với mỗi một hình tam giác và một hình vuông ở phía trên và phía dưới đường phân giác. Tiếp theo, hai hoặc nhiều phổ tương ứng với các khoảng cách được xếp hạng cao nhất, cần được loại bỏ ra khỏi quần thể phổ được làm sạch thu được ở bước b7), được gán trọng số bằng hệ số bằng 0,5, khi chúng có mặt trong cùng mặt phẳng được tạo ra bởi đường phân giác và đường thẳng đứng.

Do đó, theo một phương án ưu tiên khác, quy trình theo sáng chế còn bao gồm các dấu hiệu sau:

- ở bước c8) đường thẳng đứng được vẽ đồ thị qua đường phân giác ở một nửa chiều dài của đường phân giác này, để chia mặt phẳng giữa các trục của đồ thị phân tán thành bốn mặt phẳng, và

- ở bước c10) các phổ tương ứng với ít nhất hai khoảng cách được xếp hạng cao nhất được gán trọng số bằng hệ số bằng 0,5, khi chúng có mặt trong cùng mặt phẳng.

Về nguyên tắc, cách tiếp cận được mô tả nêu trên có thể được thực hiện cho đến khi hàm dự đoán và đường phân giác tương đẳng. Trong thực tế, một cách tiếp cận sẽ lặp lại quy trình thu được ở các bước c7) đến c11) sẽ được lặp lại cho đến khi độ hội tụ của hàm dự đoán với đường phân giác không được cải thiện hơn nữa, hoặc nói cách khác cho đến khi đồ thị của hàm dự đoán được hiệu chỉnh thu được dưới dạng đường hồi quy càng gần càng tốt với đường phân giác.

Theo một phương án ưu tiên khác, quy trình theo sáng chế còn bao gồm các bước sau: c12) lặp lại các bước c7) đến c11) với hàm dự đoán được hiệu chỉnh và quần thể phổ được làm sạch thu được từ bước hiệu chỉnh trước đó cho đến khi đồ thị của hàm dự đoán được hiệu chỉnh thu được dưới dạng đường hồi quy càng gần càng tốt với đường phân giác.

Có thể có lợi để xử lý các phổ thu được ở bước a) bằng bước xử lý dữ liệu để tạo điều kiện thuận lợi cho phân tích tiếp theo của các phổ.

Theo một phương án khác, phương pháp theo sáng chế còn bao gồm, trước bước b), bước a3) xử lý các phổ hồng ngoại thu được ở bước a) bằng bước xử lý dữ liệu, trong đó bước xử lý dữ liệu này là một hoặc nhiều bước được chọn từ nhóm bao gồm bước làm nhiễu, bước hiệu chỉnh phân tán nhân lên, bước biến đổi đại lượng ngẫu nhiên định danh chuẩn, bước loại trừ xu thế, bước lấy đạo hàm của phổ, và bước chuẩn hóa trực tiếp từng mảng liên tục.

Bản thân bước thu nhận các phổ được lấy theo thời gian tạo ra một số dạng của biến thiên ngẫu nhiên hoặc nhiễu. Do đó, bước làm nhẵn là phương pháp thích hợp để loại bỏ nhiễu và hiển thị tốt hơn tín hiệu của các quy trình nhân quả cơ bản. Theo sáng chế, tốt hơn nếu bước làm nhẵn là bước làm nhẵn trung bình động, bước làm nhẵn đa thức bằng bộ lọc Savitzky-Golar và/hoặc bước làm nhẵn đa thức tỷ lệ tín hiệu được kéo dài và nhiễu.

Bước tính toán trung bình động bao gồm bước tạo ra các chuỗi mới trong đó các giá trị bao gồm giá trị trung bình của quan sát thô trong chuỗi thời gian gốc. Trung bình động yêu cầu chỉ rõ kích cỡ cửa sổ được gọi là chiều rộng cửa sổ. Bước này xác định số lượng của quan sát thô được sử dụng để tính toán giá trị trung bình động. Phần “động” trong trung bình động liên quan đến thực tế là cửa sổ được xác định bằng chiều rộng cửa sổ trượt dọc theo các chuỗi thời gian để tính toán các giá trị trung bình trong các chuỗi mới.

Bộ lọc Savitzky-Golar là bộ lọc kỹ thuật số có thể được áp dụng cho tập điểm dữ liệu kỹ thuật số để làm nhẵn dữ liệu, tức là để tăng độ chính xác của dữ liệu mà không làm biến dạng xu hướng tín hiệu. Điều này đạt được trong quy trình được gọi là tích chập, bằng cách gán các tập con liên tiếp của các điểm dữ liệu liên kế với đa thức bậc thấp bằng phương pháp tuyến tính bình phương cực tiểu. Khi các điểm dữ liệu cách đều nhau, giải phân tích cho các phương trình bình phương cực tiểu có thể được phát hiện, ở dạng tập đơn của “hệ số tích chập” có thể được áp dụng cho toàn bộ các tập con dữ liệu, để thu được các giá trị ước lượng của tín hiệu được làm nhẵn, (hoặc các đạo hàm của tín hiệu được làm nhẵn) ở điểm tâm của mỗi tập con.

Bước loại trừ xu thế là loại hiệu chỉnh đường nền được thực hiện thông qua phép trừ mô hình tuyến tính hoặc đa thức của đường nền từ phổ ban đầu để loại bỏ biến đổi đường nền nghiêng, thường được phát hiện trong các phổ phản xạ hồng ngoại gần của các mẫu bột.

Đại lượng ngẫu nhiên định danh chuẩn (SNV) là phương pháp xử lý sơ bộ khác thường được sử dụng do thuật toán đơn giản và hiệu quả của nó trong việc hiệu chỉnh tán xạ. Phương pháp này được sử dụng trên các phổ trong đó các thay đổi đường nền và chiều dài đường dẫn gây ra các khác biệt giữa các phổ giống hệt nhau khác.

Bước hiệu chỉnh phân tán nhân lên (MSC) được thực hiện bằng cách hồi quy phổ đo được dựa trên phổ tham chiếu, sau đó hiệu chỉnh phổ đo được bằng cách sử dụng độ dốc và đoạn bị chặn của mô hình tuyến tính này. Phương pháp xử lý sơ bộ này đã được

kiểm chứng là hữu hiệu để giảm thiểu các giá trị bù đường nền và hiệu ứng nhân. Trong nhiều trường hợp, kết quả của MSC rất giống với SNV. Tuy nhiên, nhiều chuyên gia phân tích quang phổ ưu tiên chọn SNV hơn MSC do SNV hiệu chỉnh mỗi phổ riêng biệt và không cần tệp dữ liệu đầy đủ. Phương pháp xử lý sơ bộ hiệu chỉnh tín hiệu số nhân mở rộng cho phép phân tách các hiệu ứng tán xạ ánh sáng vật lý từ các hiệu ứng hấp thụ hóa học trong các phổ từ bột hoặc dung dịch đục. Phương pháp trên cơ sở mô hình đặc biệt hữu ích để giảm thiểu biến đổi tán xạ ánh sáng phụ thuộc vào bước sóng. Sau khi xử lý sơ bộ, các phổ được hiệu chỉnh trở nên không nhạy với các biến đổi tán xạ ánh sáng và đáp ứng tuyến tính với nồng độ chất phân tích.

Các trường hợp có thể phát sinh, trong đó vị trí của pic tín hiệu trong phổ không thể định vị do các cực đại và các cực tiểu của các pic riêng biệt không thể được xác định rõ ràng do trùng lặp. Bước định vị dễ dàng hơn của các pic riêng biệt trong phổ là có thể, khi các cực tiểu và các cực đại của các pic có thể xác định dễ dàng hơn. Bước lấy đạo hàm bậc nhất của phổ tạo điều kiện thuận lợi cho việc xác định các pic trong phổ do bước này tạo ra điểm về không của các cực đại pic hoặc các cực tiểu pic. Bước lấy đạo hàm bậc hai tạo ra cực tiểu pic ở chính xác vị trí đó, trong đó cực đại pic nằm trong phổ ban đầu và ngược lại. Bước lấy đạo hàm bậc nhất hoặc đạo hàm bậc hai của phổ cũng tạo điều kiện thuận lợi cho việc xác định giá trị ngoại lệ trong quần thể phổ của nguyên liệu thức ăn chăn nuôi và/hoặc thức ăn chăn nuôi đã biết. Do đó, theo sáng chế tốt hơn nếu lấy đạo hàm bậc nhất và/hoặc đạo hàm bậc hai của phổ.

Bước chuẩn hóa trực tiếp từng mảng liên tục (CPDS) được sử dụng xác định biến đổi trong các phổ thu được từ các yếu tố bên ngoài liên tục, như nhiệt độ. Bước thứ nhất trong các thuật toán PDS và CPDS là để thiết kế mô hình PLS giữa các phổ, $X(t_{ref})$ là bậc của $m \times n$, và ma trận tương ứng, Y , trong nhiệt độ cụ thể t_{ref} . Theo thuật toán chuẩn hóa trực tiếp từng mảng (PDS), ma trận được tính toán mà biến đổi các đại lượng đo phổ được ghi ở các giá trị nhiệt độ khác với ở t_{ref} :

$$X(t_{ref}) = X(t_k) \times Q(t_k), (k = 1, 2, \dots, k; t_k \neq t_{ref})$$

trong đó $Q(t_k)$ là $m \times m$ ma trận biến đổi được phân lớp được tạo ra bằng cách hồi quy tuyến tính cột thứ j của $X(t_{ref})$ trên ma trận con được lấy từ $X(t_k)$ (tức là các phổ trong $X(t_k)$ tương ứng với cửa sổ của các bước sóng từ $j - w$ đến $j + k$), bằng cách sử

dụng PLS, trong đó kích cỡ của số, $2w + 1$, được xác định bằng thông số điều chỉnh w . $Q(t_k)$ được tạo ra bằng cách thiết lập các vector hồi quy trên các khối tương ứng. Quy trình này được lặp lại đối với mỗi giá trị nhiệt độ t_k , ($k = 1, 2, \dots, k; t_k \neq t_{ref}$), và K ma trận biến đổi được tạo ra với lưu ý rằng $Q(t_k)$ là ma trận đơn vị. Để xác định tác dụng của các giá trị nhiệt độ liên tục, hàm đa thức được gán dựa trên độ chênh lệch nhiệt độ Δt_k , trong đó $\Delta t_k = t_k - t_{ref}$, với các phần tử khác 0 của ma trận được phân lớp $Q(t_k)$:

$$q_{ij}(t_k) = a_{ij} + b_{ij}\Delta t_k + c_{ij}\Delta t_k^2 + e_{ij}$$

trong đó $q_{ij}(t_k)$ là phần tử của $Q(t_k)$ đối với hàng thứ i và cột thứ j . Khi các thông số củahàm đa thức đã được ước lượng, ma trận biến đổi đối với các phổ được đo ở giá trị nhiệt độ chưa quan sát được, t_{test} , trong giai đoạn thử nghiệm có thể được tính toán bằng cách gán công thức nêu trên đối với $q_{ij}(t_k)$. Do đó, ảnh hưởng của nhiệt độ có thể được loại bỏ bằng cách biến đổi các ma trận phổ như khi được đo ở nhiệt độ tham chiếu:

$$X(t_{ref}|t_{test}) = X(t_{test}) \times Q(t_{test})$$

Sau đó các đáp ứng có thể được dự đoán bằng mô hình PLS được xây dựng trong điều kiện nhiệt độ tham chiếu.

Tốt hơn nữa, các phổ hồng ngoại thu được ở bước a) được xử lý bằng bước biến đổi đại lượng ngẫu nhiên định danh chuẩn, bước loại trừ xu thế và bước lấy đạo hàm bậc nhất của phổ, tốt hơn nữa, theo thứ tự cụ thể.

Điều này không chỉ tạo điều kiện thuận lợi mà còn quan trọng hơn là cải thiện chất lượng dự đoán của hàm dự đoán được tạo ra bằng phương pháp theo sáng chế khi chỉ các tín hiệu, mà ở các bước sóng cách đều nhau từ một tín hiệu khác trong phổ, được xem xét. Cách tiếp cận này làm giảm đáng kể nguy cơ thất thoát các tín hiệu và thông tin liên quan bất kỳ trong phổ. Điều này đặc biệt liên quan đến các dãy nồng độ của các hỗn hợp phức tạp chứa nhiều hơn chỉ một thành phần có hoạt tính hồng ngoại, trong đó các cường độ tín hiệu tăng đáng kể với nồng độ của mỗi thành phần. Do đó, các tín hiệu của thành phần có nồng độ thấp trong hỗn hợp phức tạp này có thể bị xếp chồng một phần bởi các tín hiệu của một thành phần khác có nồng độ cao hơn. Trong trường hợp này, the neglection hoặc thất thoát chỉ tín hiệu đơn trong phổ có thể dẫn đến mối tương quan dữ liệu sai và/hoặc tạo ra hàm dự đoán sai, do đó các kết dự đoán sai.

Do đó, theo một phương án khác của phương pháp theo sáng chế, các tín hiệu có các khoảng cách bước sóng cách đều nhau từ một tín hiệu khác trong phổ được xem xét ở bước b) và/hoặc c).

Trong các trường hợp trong đó bước xử lý dữ liệu dẫn đến các khác biệt trong các thang bước sóng của phổ trước và sau bước xử lý dữ liệu, có lợi để hiệu chỉnh toàn bộ phổ. Điều này đặc biệt liên quan đến vị trí chính xác của các bước sóng của các điểm vô cùng, tức là các tín hiệu có cực đại hoặc cực tiểu tương đối của độ hấp thụ. Do đó tốt hơn nếu xử lý phổ có các khoảng cách bước sóng không cách đều nhau bằng loại hiệu chỉnh để thu được phổ chính xác có các khoảng cách bước sóng cách đều nhau. Các bước sóng chính xác của các điểm vô cùng, tức là các điểm của cực đại tương đối của các độ hấp thụ và các cực tiểu tương đối của các độ hấp thụ, là cơ sở hữu hiệu để hiệu chỉnh toàn bộ phổ. Trước tiên, vị trí của các điểm vô cùng trong một hoặc nhiều phổ được dịch chuyển từ các bước sóng được dịch chuyển, tức là không hiệu chỉnh, đến các bước sóng được hiệu chỉnh, trong đó các tín hiệu tương ứng được mong muốn. Ví dụ, các tín hiệu cụ thể của các nhóm chức trong hợp chất là đã biết để được phát hiện ở các bước sóng cụ thể trong phổ hồng ngoại. Trong bước tiếp theo, đồ thị được vẽ qua các điểm vô cùng ở các bước sóng được hiệu chỉnh và trong đó thích hợp, mặc dù các điểm khác của phổ, có các vị trí không bị ảnh hưởng bởi độ dịch chuyển của các bước sóng. Bước này có thể được thực hiện bằng phép nội suy. Tuy nhiên, khó hoặc thậm chí không thực sự khả thi để vẽ chỉ một đồ thị duy nhất qua toàn các điểm vô cùng trong phổ cần hiệu chỉnh. Hơn nữa, các kết quả tốt hơn đạt được khi một số đa thức được sử dụng để nối các điểm liền kề trong phổ cần hiệu chỉnh, do chúng kết hợp tối ưu với một đồ thị duy nhất. Bước này có thể được thực hiện bằng phép nội suy hàm nối trực bậc ba. Phép nội suy hàm nối trực bậc ba có thể được thực hiện dưới dạng dải linh hoạt được uốn cong để đi qua mỗi các điểm vô cùng trong phổ có biểu diễn đồ thị cần được nội suy. Dưới góc độ toán học, dải này có thể được mô tả bằng một loạt các đa thức bậc ba, phải đáp ứng các yêu cầu là i) hàm nối trực thu được phải đi qua toàn bộ các giá trị hàm số, khi xem xét biểu diễn đồ thị của phổ dưới dạng hàm số, hoặc hàm xác định từng đoạn, ii) đạo hàm bậc nhất và đạo hàm bậc hai của các đa thức bậc ba là liên tục, và iii) độ cong buộc phải bằng 0 ở các điểm cuối của khoảng đoạn trong phổ. Khi các bước sóng được hiệu chỉnh của các điểm vô cùng là đã biết, các đa thức bậc ba có thể dễ dàng được tính toán bằng chương trình toán học thích hợp bất kỳ có bán trên thị trường. Do đó, các yêu cầu thấp về khả năng tính toán là một ưu điểm khác của phép nội suy hàm nối trực

bậc ba. Do đó, phép nội suy hàm nối trục bậc ba hoặc bước nội suy hàm nối trục bậc ba rất hữu ích để nội suy dữ liệu trong phổ to các bước sóng mới và tạo ra tính liên tục trong phổ được hiệu chỉnh.

Theo một phương án khác, phương pháp theo sáng chế còn bao gồm, trước bước b), bước a3) xử lý phổ có các khoảng cách bước sóng không cách đều nhau bằng bước nội suy hàm nối trục bậc ba để tạo ra các giá trị phổ có các khoảng cách bước sóng cách đều nhau.

Tốt hơn nữa, bước a3) được thực hiện sau bước a2) và trước bước b).

Phương pháp theo sáng chế không bị giới hạn bất kỳ liên quan đến số lượng của các hàm dự đoán cần được tạo ra. Khi số lượng cụ thể của các quần thể khác nhau của các phổ được xác định ở bước a), số lượng tương ứng của các hàm dự đoán đối với mỗi quần thể sẽ được tạo ra. Theo cách khác, khi quần thể phổ hồng ngoại thu được ở bước a) bao gồm các phổ hồng ngoại của các nhóm khác nhau của các mẫu, phương pháp theo sáng chế sẽ tạo ra hàm dự đoán tương ứng đối với mỗi nhóm này của các mẫu.

Trong phương pháp theo sáng chế phổ hồng ngoại của mẫu có nguồn gốc và/hoặc thành phần chưa biết hoặc có cùng nguồn gốc và/hoặc thành phần như mẫu ở bước a) được xác định ở bước d) và sau đó, giá trị thuộc tính quan tâm được dự đoán từ phổ này bằng hàm dự đoán thu được ở bước c) đối với mẫu của phổ thu được ở bước d). Vì mục đích này, cần chọn lọc hàm dự đoán thích hợp phù hợp với phổ mẫu truy vấn. Về nguyên tắc, hàm dự đoán có thể được chọn lọc trên cơ sở thông tin sẵn có đối với mẫu có phổ hồng ngoại được xác định ở bước d), ví dụ thông tin được tạo ra trên túi thức ăn dùng cho vật nuôi, thức ăn chăn nuôi v.v.. Tuy nhiên, thông tin này có thể không chính xác, do đó dẫn đến chọn lọc hàm dự đoán sai, hoặc người sử dụng phương pháp có thể chọn lọc hàm dự đoán theo quyết định của riêng mình, tuy nhiên hàm dự đoán này có thể là hàm dự đoán sai liên quan đến phổ mẫu truy vấn.

Do đó, tốt hơn nếu phương pháp theo sáng chế chọn lọc hàm dự đoán phù hợp nhất với phổ hồng ngoại của mẫu truy vấn. Cách tiếp cận thích hợp là để xác định độ tương tự giữa phổ mẫu được xác định ở bước d) và quần thể phổ được làm sạch bất kỳ thu được ở bước b7) hoặc trọng tâm của quần thể phổ được làm sạch này. Bước thứ nhất của Cách tiếp cận này bao gồm biến đổi cường độ hấp thụ của các bước sóng hoặc các số sóng trong phổ thu được ở bước d) thành vector truy vấn và so sánh vector truy vấn này với vector cơ sở dữ liệu thu được từ biến đổi cường độ hấp thụ của các bước sóng hoặc các số sóng trong

quần thể phổ được làm sạch thu được ở bước b7) thành tập vector cơ sở dữ liệu hoặc từ biến đổi cường độ hấp thụ của các bước sóng hoặc các số sóng của trọng tâm của quần thể phổ được làm sạch thu được ở bước b7) thành vector cơ sở dữ liệu. Tốt hơn nếu so sánh vector truy vấn với mỗi vector cơ sở dữ liệu của tập vector cơ sở dữ liệu. Theo nghĩa rộng nhất của nó, vector là đối tượng hình học có kích cỡ (hoặc chiều dài) và hướng. Trong hệ tọa độ Đề các, vector có thể được biểu diễn bằng cách xác định các tọa độ của điểm đầu và điểm cuối của nó. Do đó, vector thích hợp để biểu diễn cường độ độ hấp thụ ở bước sóng hoặc số sóng cụ thể trong phổ hồng ngoại gần hai chiều. Ngoài ra, vector không chỉ giới hạn ở hệ hai chiều. Hơn nữa, vector có thể mô tả không gian nhiều chiều, như phổ hồng ngoại gần có nhiều cường độ hấp thụ ở nhiều bước sóng hoặc số sóng khác nhau. Trong trường hợp này, mỗi chiều của vector này tương ứng với cường độ độ hấp thụ duy nhất ở bước sóng hoặc số sóng cụ thể. Bước tiếp theo là tính toán giá trị đo độ tương tự và/hoặc giá trị đo khoảng cách giữa vector truy vấn và vector cơ sở dữ liệu hoặc mỗi vector của tập vector cơ sở dữ liệu và xác định hạng giá trị đo độ tương tự và/hoặc giá trị đo khoảng cách thu được có giá trị tốt nhất ở đỉnh của hạng, tức là giá trị đo độ tương tự với giá trị cao nhất ở đỉnh hoặc giá trị đo khoảng cách có giá trị thấp nhất ở đỉnh. Hàm dự đoán tương ứng với quần thể phổ được làm sạch từ đó vector cơ sở dữ liệu có độ tương tự cao nhất ở bước d4) được chọn để dự đoán ở bước e).

Do đó, tốt hơn nếu, bước d) của phương pháp theo sáng chế còn bao gồm các bước sau:

d1) biến đổi cường độ hấp thụ của các bước sóng hoặc các số sóng trong phổ thu được ở bước d) thành vector truy vấn,

d2) biến đổi cường độ hấp thụ của các bước sóng hoặc các số sóng in quần thể phổ được làm sạch thu được ở bước b7) thành tập vector cơ sở dữ liệu hoặc biến đổi cường độ hấp thụ của các bước sóng hoặc các số sóng của trọng tâm của quần thể phổ được làm sạch thu được ở bước b7) thành vector cơ sở dữ liệu,

d3) tính toán giá trị đo độ tương tự và/hoặc giá trị đo khoảng cách giữa vector truy vấn thu được ở bước d1) và vector cơ sở dữ liệu thu được ở bước d2) để thu được giá trị độ tương tự đối với vector truy vấn với vector cơ sở dữ liệu,

d4) xếp hạng các giá trị độ tương tự thu được ở bước d3) theo thứ tự giảm dần, khi giá trị đo độ tương tự được tính toán ở bước d3) hoặc theo thứ tự tăng dần, khi giá trị đo

khoảng cách được tính toán ở bước d3), trong đó trong trường hợp bất kỳ vector cơ sở dữ liệu được xếp hạng cao nhất có độ tương tự cao nhất với vector truy vấn, và

d5) selecting hàm dự đoán tương ứng với quần thể phổ được làm sạch từ which vector cơ sở dữ liệu có độ tương tự cao nhất ở bước d4) để dự đoán ở bước e).

Tốt hơn nếu biến đổi cường độ hấp thụ ở các bước sóng hoặc các số sóng cách đều nhau trong phổ thu được ở bước d) và trong quần thể phổ được làm sạch thu được ở bước b7) và cụ thể ở cùng các bước sóng hoặc các số sóng cách đều nhau trong phổ thu được ở bước d) và trong quần thể phổ được làm sạch thu được ở bước b7) thành vector. Điều này cho phép phân tích độ tương tự tốt nhất có thể. Tốt hơn nếu, khoảng cách giữa các bước sóng ở bước d1) và/hoặc d2) của phương pháp theo sáng chế nằm trong khoảng từ 0,1 +/- 10% đến 10 +/- 10% nm, nằm trong khoảng từ 0,1 +/- 10% đến 5 +/- 10% nm, hoặc nằm trong khoảng từ 0,1 +/- 10% đến 2 +/- 10% nm. Theo đó, khoảng cách giữa các số sóng ở bước d1) và/hoặc d2) của phương pháp theo sáng chế nằm trong khoảng từ 108 +/- 10% đến 106 10 +/- 10%, nằm trong khoảng từ 108 +/- 10 đến 5 * 106 +/- 10% nm, hoặc nằm trong khoảng từ 108 +/- 10% đến 2 * 106 +/- 10% nm. Thuật ngữ “ +/- 10%” được sử dụng liên quan đến các giá trị được đề cập rõ ràng để chỉ ra rằng độ lệch từ các giá trị được đề cập rõ ràng này vẫn nằm trong phạm vi của sáng chế, đã xác định được rằng về cơ bản chúng tạo ra các hiệu quả của sáng chế.

Phân tích độ tương tự ở bước d3) không bị giới hạn ở việc sử dụng khoảng cách hoặc giá trị đo độ tương tự cụ thể. Về nguyên tắc, phân tích độ tương tự là phân tích để tra cứu vector lân cận gần nhất. Do đó, giá trị đo độ tương tự bất kỳ thích hợp để xác định vector lân cận gần nhất với vector truy vấn có thể được sử dụng trong ở bước d3).

Độ tương tự cosin là ví dụ về giá trị đo độ tương tự thích hợp để sử dụng theo sáng chế. Độ tương tự cosin cho phép tính toán độ tương tự giữa hai vector cực kỳ nhanh chóng và đồng thời với độ chính xác cao. Độ tương tự cosin $C_{A,B}$ giữa hai quan sát, tức là hai vector của các phổ, có thể được tính toán bằng công thức

$$C_{A,B} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

trong đó A_i và B_i là các thành phần của các vector A và B , trong số đó một vector là vector truy vấn và vector còn lại là vector cơ sở dữ liệu, và n là số lượng của các vector được

xem xét. Các giá trị cho giá trị đo độ tương tự có thể nằm trong khoảng từ -1 có nghĩa là hoàn toàn trái ngược nhau, qua 0, thể hiện tính trực giao (bất tương quan) đến +1, có nghĩa là tương đồng, với các giá trị ở giữa thể hiện độ tương tự hoặc độ không tương tự trung gian.

Khoảng cách Euclid là ví dụ về giá trị đo khoảng cách thích hợp để sử dụng theo sáng chế Khoảng cách Euclid cho phép tính toán độ không tương tự giữa hai vector cực kỳ nhanh chóng và đồng thời với độ chính xác cao. Khi cần, giá trị đo khoảng cách thu được có thể dễ dàng được tính toán thành giá trị đo độ tương tự để sử dụng trong phương pháp theo sáng chế. Khoảng cách Euclid $E_{A,B}$ giữa hai quan sát, tức là hai vector của các phổ, có thể được tính toán bằng công thức

$$E_{A,B} = \sqrt{\sum_{i=1}^n (A_i^2 - B_i^2)^2}$$

trong đó A_i và B_i là các thành phần của các vector A và B , trong số đó một vector là vector truy vấn và vector còn lại là vector cơ sở dữ liệu, và n là số lượng của các vector được xem xét. Các giá trị cho giá trị đo khoảng cách có thể nằm trong khoảng từ +1 có nghĩa là hoàn toàn trái ngược nhau, qua 0, thể hiện tính trực giao (bất tương quan) đến -1, có nghĩa là tương đồng, với các giá trị ở giữa thể hiện độ tương tự hoặc độ không tương tự trung gian.

Sáng chế cũng đề cập đến hệ thống dự đoán giá trị thuộc tính quan tâm của nguyên liệu, bao gồm bộ xử lý, được làm thích ứng để thực hiện phương pháp dự đoán giá trị thuộc tính quan tâm của nguyên liệu được thực hiện trên máy tính theo sáng chế và/hoặc các phương án bất kỳ của nó.

Một hoặc nhiều hàm dự đoán được tạo ra trong phương pháp theo sáng chế được lưu trữ trong bộ xử lý của hệ thống. Điều này cho phép liên lạc giữa hệ thống theo sáng chế và một hoặc nhiều thiết bị quang phổ hồng ngoại ở cùng vị trí như hệ thống này và/hoặc ở vị trí khác.

Theo một phương án của hệ thống theo sáng chế bộ xử lý tạo ra mạng có một hoặc nhiều thiết bị quang phổ hồng ngoại.

YÊU CẦU BẢO HỘ

1. Phương pháp dự đoán giá trị thuộc tính quan tâm của nguyên liệu được thực hiện trên máy tính, bao gồm các bước sau:

a) xác định quần thể phổ hồng ngoại của các mẫu, trong đó các phổ này tạo ra $m \times n$ ma trận dữ liệu đầu vào X , với m là số lượng mẫu trong các hàng và n là các điểm dữ liệu trong các cột,

b) loại bỏ các giá trị ngoại lệ phổ ra khỏi quần thể phổ thu được ở bước a), bao gồm các bước sau:

b1) thu nhận các thành phần chính bằng cách xử lý ma trận X bằng phân tích thành phần chính,

b2) tạo ra ma trận chéo Σ , chứa các giá trị suy biến σ_m của ma trận X , và ma trận hệ số tải V , từ ma trận dữ liệu đầu vào X ,

b3) tính toán điểm số x_m đối với mỗi phổ bằng cách nhân mỗi điểm dữ liệu của ma trận dữ liệu đầu vào X với các hệ số tải đối với mỗi thành phần thu được ở bước b2), tạo ra giá trị trung bình của mỗi cột của ma trận X để tạo ra các giá trị $B_{0,m}$, và tính toán chỉ số chấm điểm si bằng công thức

$$si = \sum_i \frac{(x_m - B_{0,m})}{\sigma_m}$$

b4) xác định số lượng của các thành phần N_c có các giá trị giằng dẫn đến sự hội tụ trong hồi quy của X trên các điểm số ít nhất bằng 99%, và tính toán giá trị giới hạn đo khoảng cách D_i đối với mọi phổ thu được ở bước a), bằng công thức

$$D_i = N_c \times \sqrt{\frac{(si^2 \times N_c)}{(N_c - 1)}}$$

b5) tính toán giá trị trung bình đối với toàn bộ các điểm số của mỗi thành phần chính của mỗi phổ thu được ở bước a) và tính toán giá trị đo khoảng cách giữa giá trị trung bình này và mỗi điểm số của mỗi thành phần chính,

b6) xem xét phổ mẫu là giá trị ngoại lệ phổ khi giá trị của giá trị đo khoảng cách đối với điểm số của thành phần chính thu được ở bước b5) lớn hơn giá trị giới hạn đo khoảng cách thu được ở bước b4),

b7) loại bỏ giá trị ngoại lệ phổ thu được ở bước b6) ra khỏi quần thể phổ thu được ở bước a) để thu được quần thể phổ được làm sạch,

c) tạo ra hàm dự đoán trên quần thể phổ được làm sạch thu được ở bước b7),

d) xác định phổ hồng ngoại của mẫu có nguồn gốc và/hoặc thành phần chưa biết hoặc có cùng nguồn gốc và/hoặc thành phần như mẫu ở bước a), và

e) dự đoán giá trị thuộc tính quan tâm từ phổ thu được ở bước d) bằng hàm dự đoán thu được ở bước c).

2. Phương pháp theo điểm 1, trong đó giá trị đo khoảng cách là giá trị đo khoảng cách Euclid, giá trị đo khoảng cách Pearson, giá trị đo khoảng cách Mahalanobis, hoặc giá trị đo khoảng cách thu được từ giá trị đo độ tương tự.

3. Phương pháp theo điểm 1 hoặc 2, trong đó bước b) còn bao gồm các bước sau:

b5.1) tăng giá trị giới hạn đo khoảng cách thu được ở bước b4) bằng +1,

b5.2) xác định hai giá trị đo khoảng cách thu được ở bước b5) có các giá trị cao nhất bằng cách sử dụng giá trị giới hạn đo khoảng cách thu được ở bước b5.1),

b5.3) xác định độ chênh lệch giữa các giá trị của các giá trị đo khoảng cách được xác định ở bước b5.2), và

b5.4) lặp lại các bước d5.1) đến d5.3) với giá trị giới hạn đo khoảng cách thu được ở bước b5.1) cho đến khi độ chênh lệch được xác định ở bước b5.3) ít nhất bằng 1, và giá trị cao nhất của giá trị đo khoảng cách lớn nhất bằng 8.

4. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 3, trong đó phương pháp này còn bao gồm bước a1) xác định các giá trị thuộc tính quan tâm trong mỗi mẫu thu được ở bước a) trong phân tích định lượng để thu được tập hợp dữ liệu tham chiếu.

5. Phương pháp theo điểm 4, trong đó bước tạo ra hàm dự đoán ở bước c) bao gồm bước phân tích tập hợp dữ liệu tham chiếu thu được ở bước a1) và quần thể phổ được làm sạch thu được ở bước b7) đối với mỗi tương quan dữ liệu để thu được hàm dự đoán.

6. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 5, trong đó bước tạo ra hàm dự đoán ở bước c) bao gồm bước kiểm chứng hàm dự đoán này.

7. Phương pháp theo điểm 6, trong đó bước kiểm chứng là bước kiểm chứng chéo chia đôi, bước kiểm chứng chéo k lần và/hoặc bước kiểm chứng dựa trên dữ liệu kiểm định vòng.

8. Phương pháp theo điểm 5 hoặc 6, trong đó bước c) còn bao gồm các bước sau:

c1) vẽ đồ thị các giá trị thuộc tính được dự đoán đối với quần thể phổ được làm sạch thu được ở bước b7), và hàm dự đoán thu được ở bước c) dưới dạng đường hồi quy thành đồ thị giá trị dự đoán so với giá trị đo được để thu được đồ thị phân tán,

c2) vẽ đồ thị đường phân giác giữa các trục của đồ thị thu được ở bước c1),

c3) xác định các khoảng cách của mỗi giá trị thuộc tính dự đoán thu được ở bước c1) đến đường phân giác trong đồ thị phân tán thu được ở bước c2),

c4) tạo ra hạng tổng thể của các khoảng cách thu được ở bước c3), với các khoảng cách cao nhất ở đỉnh của hạng, và loại bỏ các phổ tương ứng với ít nhất ba khoảng cách được đánh giá ở đỉnh ra khỏi quần thể phổ được làm sạch thu được ở bước b7), và

c5) hiệu chỉnh hàm dự đoán thu được ở bước c) để bù giá trị loại bỏ của các phổ ở bước c4).

9. Phương pháp theo điểm 8, trong đó phương pháp này còn bao gồm bước c6) lặp lại các bước c1) đến c5) với hàm dự đoán được hiệu chỉnh và quần thể phổ được làm sạch thu được từ bước hiệu chỉnh trước đó cho đến khi đồ thị của hàm dự đoán được hiệu chỉnh thu được dưới dạng đường hồi quy càng gần càng tốt với đường phân giác.

10. Phương pháp theo điểm bất kỳ trong số các điểm từ 5 đến 9, trong đó phương pháp này còn bao gồm các bước sau:

c7) vẽ đồ thị các giá trị thuộc tính được dự đoán đối với quần thể phổ được làm sạch thu được ở bước b7), và hàm dự đoán thu được ở bước c) dưới dạng đường hồi quy thành đồ thị giá trị dự đoán so với giá trị đo được để thu được đồ thị phân tán,

c8) vẽ đồ thị đường phân giác giữa hai trục của đồ thị thu được ở bước c7),

c9) vẽ đồ thị khoảng tin cậy với chiều rộng xác định trước xung quanh đường phân giác thành đồ thị thu được ở bước c8),

c10) loại bỏ các phổ tương ứng với các điểm nằm ngoài khoảng tin cậy trong đồ thị thu được ở bước c9) ra khỏi quần thể phổ được làm sạch thu được ở bước b7), và

c11) hiệu chỉnh hàm dự đoán thu được ở bước c) để bù giá trị loại bỏ của các phổ ở bước c10).

11. Phương pháp theo điểm 10, trong đó phương pháp này còn bao gồm bước c12) lặp lại các bước c7) đến c11) với hàm dự đoán được hiệu chỉnh và quần thể phổ được làm sạch thu được từ bước hiệu chỉnh trước đó cho đến khi đồ thị của hàm dự đoán được hiệu chỉnh thu được dưới dạng đường hồi quy càng gần càng tốt với đường phân giác.

12. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 11, trong đó phương pháp này còn bao gồm, trước bước b), bước a2) xử lý các phổ hồng ngoại thu được ở bước a) bằng bước xử lý dữ liệu, trong đó bước xử lý dữ liệu này là một hoặc nhiều bước được chọn từ nhóm bao gồm bước làm nhẵn, bước hiệu chỉnh phân tán nhân lên, bước biến đổi đại lượng ngẫu nhiên định danh chuẩn, bước loại trừ xu thế, bước lấy đạo hàm của phổ, và bước chuẩn hóa trực tiếp từng mảng liên tục.

13. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 12, trong đó phương pháp này còn bao gồm, trước bước b), bước a3) xử lý phổ có các khoảng cách bước sóng không cách đều nhau bằng bước nội suy hàm nối trực bậc ba để tạo ra các giá trị phổ có các khoảng cách bước sóng cách đều nhau.

14. Hệ thống dự đoán giá trị thuộc tính quan tâm của nguyên liệu, bao gồm bộ xử lý, được làm thích ứng để thực hiện phương pháp dự đoán giá trị thuộc tính quan tâm của nguyên liệu được thực hiện trên máy tính theo điểm bất kỳ trong số các điểm từ 1 đến 13.

15. Hệ thống theo điểm 14, trong đó bộ xử lý tạo ra mạng có một hoặc nhiều thiết bị quang phổ hồng ngoại.