

報 文

クロマトグラムの積分方法が測定値の真度・精度へ与える影響

松田りえ子¹, 豊田 正武², 林 譲^{®1}

Trueness and precision of two integration methods in chromatography

Rieko MATSUDA¹, Masatake TOYODA² and Yuzuru HAYASHI¹¹ National Institute of Health Sciences, 1-18-1, Kami-Yoga, Setagaya-ku, Tokyo 158-8501² Jissen Women's University, 4-1-1, Osakaue, Hino-shi, Tokyo 191-8510

(Received 24 April 2003, Accepted 18 June 2003)

Two integration methods for measuring the peak areas and heights in chromatography are compared with special emphasis on trueness and precision. The most important difference of the methods is the integration region which is determined to cover the entire area of a peak. In the traditional integration method, the integration region can vary depending on the actual noise on the peak signal, but it is fixed for every noisy peak in the other method (called TOCO). For the traditional method, large biases are found and the resulting calibration lines have positive intercepts, but there are no such disadvantages for TOCO. The precision is almost comparable to both the methods. It is concluded that both methods have no significant problems in practical use. A fundamental obstacle of the traditional method, when the standard deviation of measurements is estimated, is pointed out.

Keywords : Chromatography; Precision; Trueness; Accuracy; Integration.

1 緒 言

クロマトグラフィーでは、なんらかの方法でピークの始まりと終わりの点を定め、このピーク領域から面積、高さを算出する。一般に、これはピーク積分法と呼ばれている。最もよく使われているピーク積分法は、設定されたパラメーターに基づいてピークの立ち上がりとしち下がりを実動的に検出して、ピーク領域を定め、面積・高さの測定値を計算する。以下、この方法をインテグレーター法と呼ぶ。

クロマトグラフィーのピークは本来ゆっくり立ち上がり、ゆっくり立ち下がるガウス型の曲線である。しかし、ノイズが大きい場合は、ピーク本来の立ち上がり・立ち下がりを実検するのは不可能であり、ノイズの立ち上がり・立ち下がりを実検することになる。そのため、インテグレーター法は、Fig. 1 に示すように、様々なピーク領域を示す。

す。但し、一番上のピーク領域は、151 データポイントに固定されている（以下参照）。

インテグレーター法では、次のような問題が生じると考えられる。

A: ピーク領域の始点はノイズの立ち上がりであり、終点はノイズの立ち下がりであるため、測定値は正しい値より大きく見積もられる可能性が高い。

B: ピーク領域がクロマトグラムごとに変わる。

A からは、インテグレーター法は真度の点での検討が必要になり、B からは、精度（測定値の標準偏差）を計算するときに大きな問題が生じる。A の問題点は明白であるが、B が指摘されたことは今までにはないと思える。

本文中で詳しく議論するが、ノイズはデータポイントごとに存在するため、異なった幅（データポイント数）のピーク領域での積分値から標準偏差を計算しても意味がない。以上は数学的に厳密な話であるので、実験によって、これが実際に重要な問題であるかどうかを調べる必要がある。そこで、本研究では、ノイズとガウスピークを用いたシミュレーション実験により、ピーク領域のデータポイント

¹ 国立医薬品食品衛生研究所: 158-8501 東京都世田谷区上用賀 1-18-1

² 実践女子大学: 191-8510 東京都日野市大坂上 4-1-1

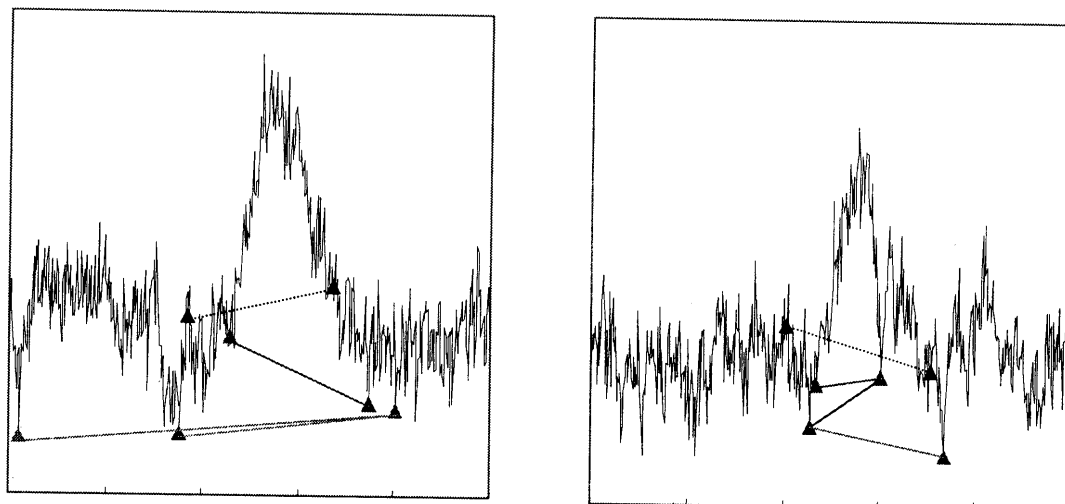


Fig. 1 Variation in the regions and zero lines for integration method and TOCO

▲: the both sides of the integration regions; —: the zero lines; the uppermost —: TOCO; the other: integration method. The peak region in TOCO is fixed to be 151 data points.

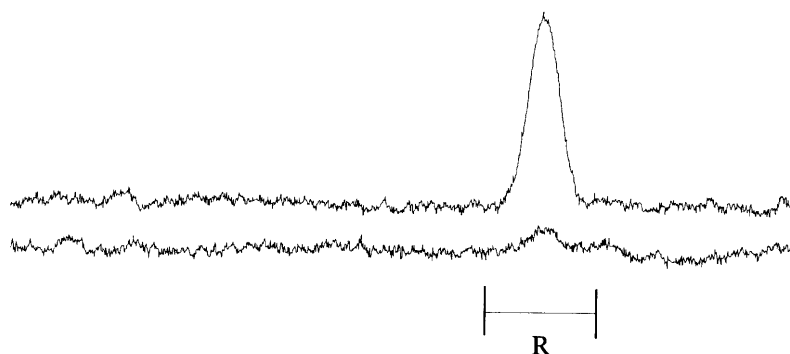


Fig. 2 Baseline noise and superimposed signal

Peak area: upper, 5000; lower, 500. Letter, R, denotes the integration region of 191 data points

ト数を固定して測定値を得る方法 (TOCO 法) とインテグレーター法を比較する。クロマトグラフィーの真度・精度を研究した論文は多い^{1)~11)}。しかし、インテグレーター法の真度・精度を他の方法と比べた報告は見当たらない。

2 実験方法

合成ノイズの上に面積既知のピークを重ねたクロマトグラムを合成し、ピーク面積・高さを測定した。

合成ノイズでは、ホワイトノイズの SD は 1, マルコフ過程の SD は 0.5, 自己相関係数は 0.95 とした。ピークはガウス型であり、面積 (高さ) は, 500 (7.98), 750 (11.97), 1500 (23.94), 2500 (39.89), 5000 (79.79) とした。Fig. 2 に合成したシグナル+ノイズの例を挙げる。

ピーク積分は 2 つのソフトで行い、測定値の真度・精度を比べた。測定に用いたソフトは、EGChrom Elite

(GL Science 製) と TOCO ver.2 であった。なお、TOCO は FUMI 理論研究会 (<http://www8.plala.or.jp/fumitheory/>) で配布している研究用ソフトである。どちらのソフトもピークの面積と高さの測定のためにパラメーターを設定する必要がある。

EGChrom (以下、インテグレーター法と呼ぶ) では、2 つのパラメーター (width と threshold) を設定する。ここでは、width = 20, 30, 40 とし、threshold = 1E + 06, 1E + 07 とした。

TOCO (以下、TOCO 法と呼ぶ) で設定するパラメーターは、ピークの範囲とピーク頂点の位置である。ピーク領域は、151, 191 データポイントであり、頂点の位置はピーク領域の中央とした。

ガウス型シグナルを固定し、ノイズを加えてクロマトグラムを 100 個合成する。これらのクロマトグラムに対してピーク積分を行う。それぞれの積分法のパラメーターは

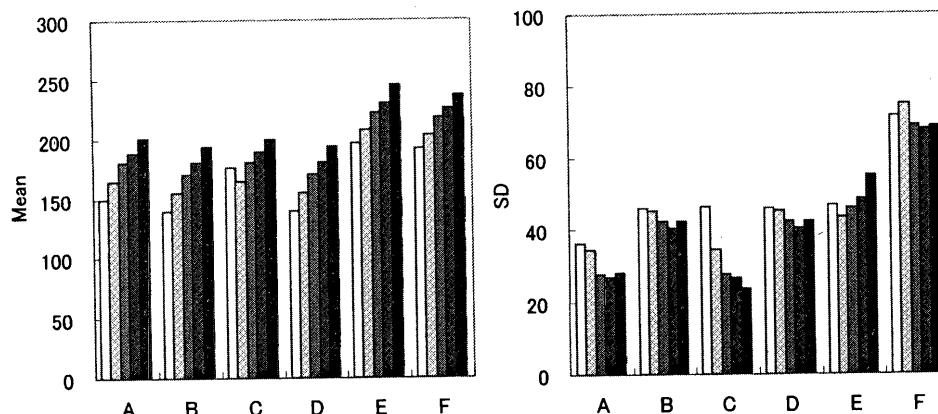


Fig. 3 Mean (left) and SD (right) of the number of data points used for integration method

A (width, threshold): 20, 10^6 ; B: 20, 10^7 ; C: 30, 10^6 ; D: 20, 10^7 ; E: 40, 10^6 ; F: 40, 10^7 . Bars denotes peak areas (from left to right): 500; 750; 1500; 2500; 5000

固定し, 100 個の測定値 (面積と高さ) を得る. これらから平均値と標準偏差を計算する. パラメーターの設定又はピークの大きさを変えて, 同様な手順を踏んで平均値と標準偏差を計算する.

TOCO では, 合成したシグナル+ノイズの値をそのまま読み込み, 高さと面積の測定値を計算するため, 真の測定値はクロマトグラムの合成に使ったガウス型ピークの高さと面積と同じになる. しかし, EGChrom は合成クロマトグラムの数値を変換して測定値を計算する. そこで, ノイズのないシグナルを EGChrom で読み込み, 得られた測定値を真の値とした. バイアスは, ノイズがある合成クロマトグラムからの測定値と真の値の差とする.

3 結果と考察

3.1 実験 1: ピークとピーク領域

多くのクロマトグラフのベースラインノイズは $1/f$ ゆらぎを示すことが知られている. ここでは, ホワイトノイズとマルコフ過程の和で $1/f$ ゆらぎを近似する方法を用いた¹⁰⁾. 実際のベースラインノイズに似るように, ホワイトノイズとマルコフ過程のパラメーターを調節した.

Fig. 2 は, ピークとノイズを重ね合わせた合成クロマトグラムを示す. 上のピークは, 検討した最大のピークであり, 下のピークは最小のピークである. 大まかに言うと, 上のピークは定量限界くらいであり (測定値の RSD = 約 10%), 下のピークは検出限界くらいである (測定値の RSD = 約 30%). 通常, 液体クロマトグラフの注入誤差は約 0.3% RSD であり, Fig. 2 にあるようなピークでは, 測定誤差はベースラインノイズだけに起因すると言って差し支えない. つまり, 本研究では, ベースラインノイズに起因する測定誤差を考察する.

検討する 2 つの測定方法 (インテグレーター法と

TOCO 法) は, どちらもピークの始まりと終わりの点を定め, それを基準として面積, 高さを算出する. Fig. 2 に示してある領域 R は, TOCO 法で設定されたピーク領域の例 (191 データポイント) である.

インテグレーター法では, 設定したパラメーター (width と threshold) を基にピークが検出されると, そのピークの面積と高さが算出される. ピークがどこにあっても測定値を得られるが, ノイズによってピーク領域 (データポイント数) は異なる. ピークが大きくなると, ピーク領域が大きくなる傾向が観測された (Fig. 3 を参照). 例えば, width = 20 と threshold = $1\text{E} + 06$ の設定 A では, 面積 500 のピークの領域の平均は約 150 であるが, 面積 5000 のピークの領域の平均は約 200 である. 測定ごとによるピーク領域のばらつきは, ピークの大きさとは無関係にほぼ一定であり, SD = 30 ~ 50 である. 但し, 設定 F (width = 40 と threshold = $1\text{E} + 07$) では, ばらつきは非常に大きい (SD = ca. 70).

TOCO 法では, ピーク領域を前もって設定するため, ピーク領域はクロマトグラムごとに変動しない. ピーク領域にあるノイズの数が測定ごとに変化しないため, ピーク測定値に与えるノイズの影響を統計的に考察するには理にかなっている. しかし, ピークの位置が大きすぎると, ピークの自動検出が難しくなる. もちろん, 手動で調整すればよい.

3.2 実験 2: 正確さの検討

インテグレーター法は, シグナル+ノイズのもっと高い位置を高さ測定値とするため, 高さ測定値のバイアスが生じ, これは正になると予想できる. また, 面積測定においても, シグナル+ノイズの立ち上がりの点をピーク開始点とし, 立ち下がりの点をピーク終点とするため, ピークの

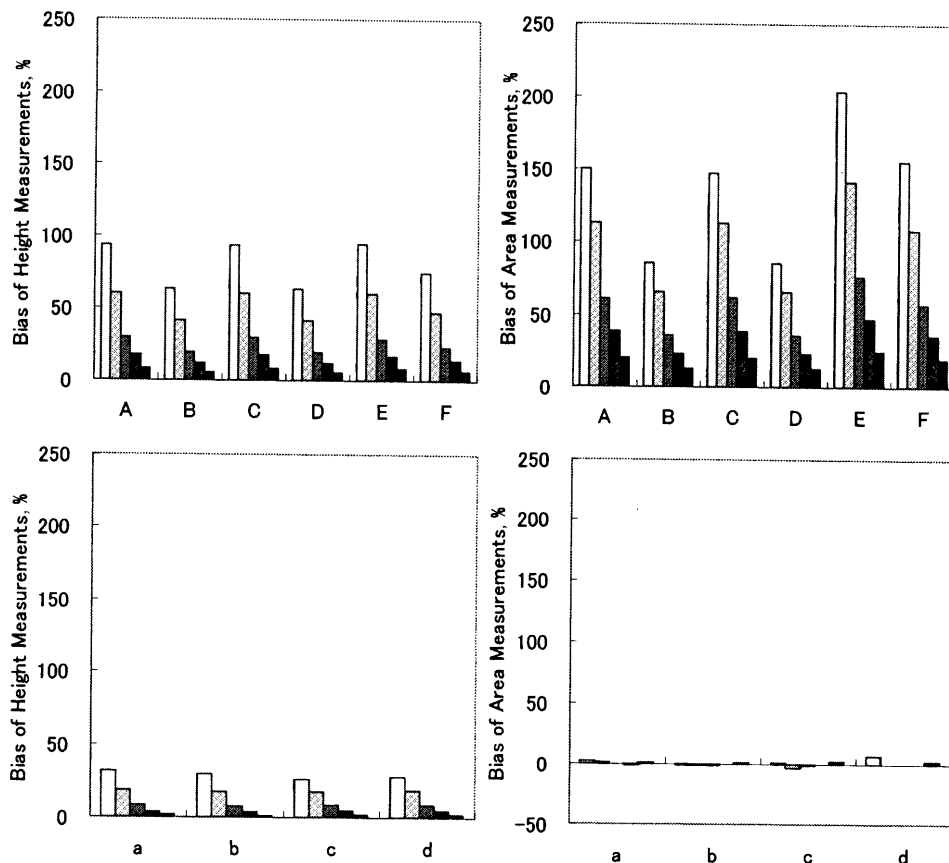


Fig. 4 Trueness of integration method (top) and TOCO (bottom)

For conditions A~F, see the legend of Figure 3. a (width baseline): 151, horizontal; b: 151, oblique; c: 191, horizontal; d: 191, oblique

両端がげたを履いているようなもので、面積測定値も実際より大きく算出されることが考えられる。

Fig. 4に示す実験結果もこの議論を裏付ける。高さ測定・面積測定共にピークが小さいほど相対的バイアスは大きい。最も大きいピーク（面積5000）では、高さ測定値のバイアスは10%以下であるが、最も小さいピーク（面積500）では、60~100%と非常に大きい。面積測定値のバイアスは更に大きい。なお、インテグレーター法では、パラメーター設定にはよらず、常に正のバイアスが生じる。

TOCO法では、上のようなバイアスを避けるために、シグナル+ノイズのもっと高い位置を検出した後、ピーク領域とピーク頂点の位置をランダムに数ポイント内で移動する設定もある。しかし、インテグレーター法と公平に比較するため、本研究ではシグナル+ノイズのもっと高い位置を高さ測定の位置とした。なお、高さ測定の位置は、ピーク領域の中心に設定してあるため、ピーク領域は高さ測定位置から逆算して決まる。ピーク領域の幅（データポイント）は固定であるので、ピーク領域は、ピークごとに平行移動すると思えばよい。高さ測定値の相対的バイアスは、最も小さいピークでも約30%であり、面積測定値の

相対的バイアスは更に小さい（10%以下）。TOCO法のバイアスはインテグレーター法よりかなり小さい。

TOCO法では、ピーク頂点の位置を基準にピーク領域を決めるため、ピークの開始点ではシグナル+ノイズが上昇することも下降することも同確率で起こる。そのため、正の誤差も負の誤差も観測される。しかし、平均的には、面積測定値のバイアスはほとんどない。

3.3 実験3: 精密さの検討

クロマトグラフィー測定の精度の一般的性質は、ピークが小さくなると測定値のRSDは大きくなることである¹⁰⁾。この傾向はFig. 5でも示されている。高さ測定でも面積測定でも、この傾向は同じであり、パラメーターの設定にもよらない。

インテグレーター法ではバイアスは生じるが、測定値のばらつきは小さく抑えられると予想できる。なぜならば、インテグレーター法はランダム変動の立ち上がりとしち下がり必ずピーク開始点と終点とするからである。

Fig. 5より、最も小さいピークに対して、TOCO法の精度は、高さ測定では約20% RSDであり、面積測定では、約40~60% RSDである。一方、対応するインテグレータ

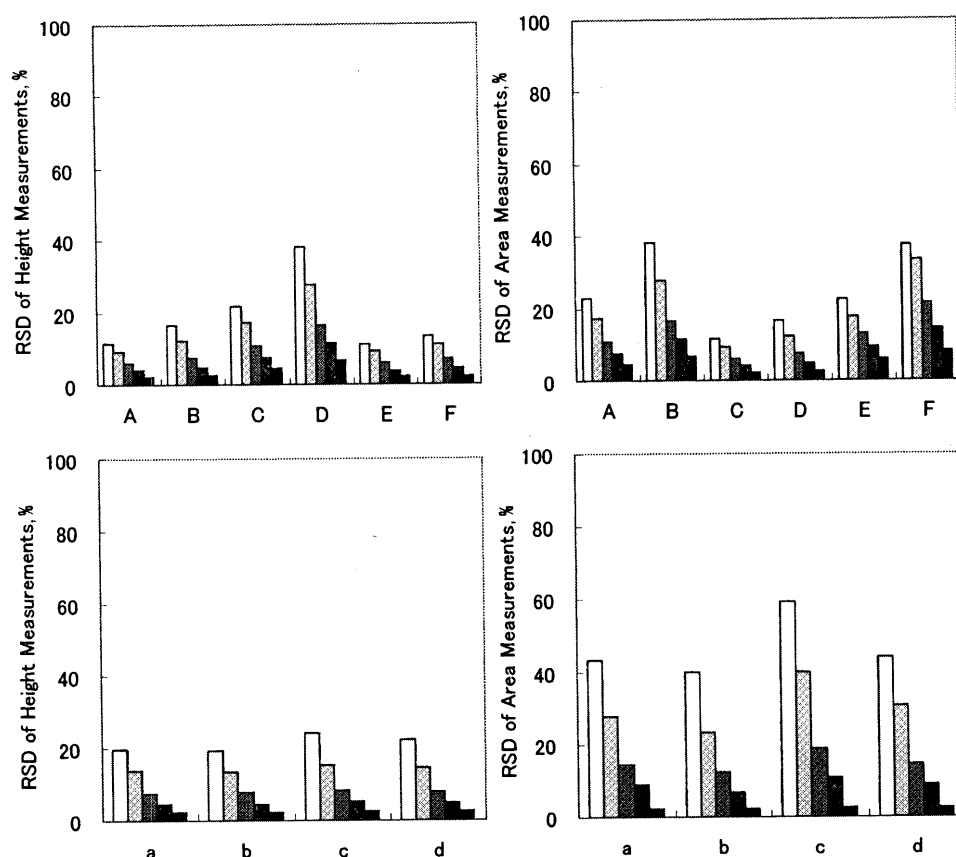


Fig. 5 Precision of integration method (top) and TOCO (bottom)
For conditions A~F and a~d, see the legends of Figure 3 and 4.

一法の精度は、高さ測定でも面積測定でも、約 10~40% RSD である。しかし、大きなピークに対しては、これらの方法にはほとんど差がないことも分かる。

パラメーターの設定によっては、インテグレーター法は TOCO 法より少し優れていることもあるが、全体的に見ると、インテグレーター法が精度において有意に優れているとは言い難い。

3.4 実験 4: 検量線の検討

Fig. 6 は、インテグレーター法と TOCO 法から求めた検量線を示す。TOCO 法はバイアスが小さいため、検量線は原点を通る直線になる。測定値の精度も問題がないため、定量分析には向いている方法であろう。

インテグレーター法にはバイアスがあるため、検量線は有限の y -切片をもつが、直線になった。但し、パラメーターの設定によって y -切片の大きさは異なる。バイアスの程度によっては検量線が曲がる可能性が考えられるが、ここで検討した条件では、どの検量線も直線となった。そのため、実用的な面に限れば、インテグレーター法も定量分析的には問題がないと考えられる。

インテグレーター法の測定値は実際のピーク高さ・面積とは異なる。これは、用いたソフト (EGChrom Elite) が

数値を変換しているためである。

3.5 インテグレーター法の問題点

クロマトグラムはデジタルで記録され、ピーク領域はデータポイント数で表される。また、低濃度測定においては、測定誤差の主な原因はベースラインノイズであり、ノイズはデータポイントごとに存在する。

その結果、インテグレーター法では、同じサンプルでもクロマトグラムごとにピーク領域 (データポイント数) が異なる。異なったピーク領域での測定値を同等に扱って (つまり、混ぜて)、標準偏差を計算するのは理論的に誤りであることを以下で議論する。

目が -3, -2, -1, 1, 2, 3 である正 6 面体のサイコロをノイズのモデルとする。どの目が出る確率も等しい ($=1/6$) と仮定すれば、目の平均は 0 であり、標準偏差は 2.16 である。

$$\left(\because \sqrt{\frac{(-3)^2 + (-2)^2 + (-1)^2 + 1^2 + 2^2 + 3^2}{6}} \approx 2.16 \right)$$

ピーク領域が 10 データポイントである場合、このサイコロを 10 回振って出た目の合計を測定誤差とする。20 データポイントのピーク領域の測定誤差は、サイコロを 20

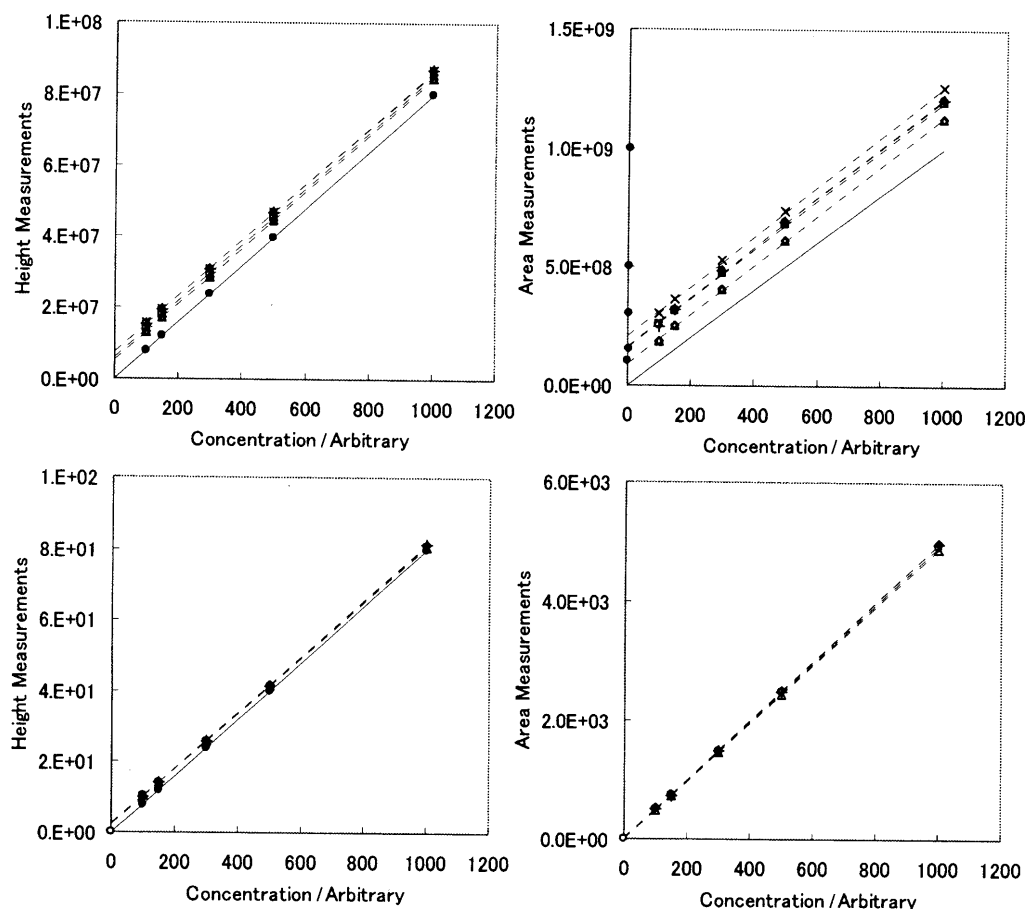


Fig. 6 Calibration lines of integration method (top) and TOCO (bottom)

Integration: $\blacklozenge$, true value; $\circ$, A; $\blacktriangle$, B; $+$, C; $\diamond$, D; $\times$, E; $\square$, F (for conditions A~F, see the legend of Figure 3). TOCO: $\circ$, a; $\blacktriangle$, b; $+$, c; $\diamond$, d (for conditions a~d, see the legend of Figure 4).

回振って出た目の合計である。どちらのピーク領域でも、測定誤差の平均は0である。

平均値とは違い、ピーク領域の幅が異なると、標準偏差は異なる。10 データポイントの領域では、サイコロの合計は-30 から+30 の間で変動し、20 データポイントの場合は、-60 から+60 の間で変動する。変動の幅が異なるので、これらの領域での測定誤差の標準偏差も当然異なる。分散の加法性より、これらの標準偏差は理論的に計算でき、それぞれ、 $2.16\sqrt{10}$ と $2.16\sqrt{20}$ となる。

真の統計量が不明な場合は、繰り返し実験から真の値を推定する。例えば、10 データポイントのピーク領域での測定誤差の標準偏差を知りたいときは、サイコロを10回投げる実験を、例えば、6回繰り返し（合計 10×6 ）、得られた6個の値から真の値 $2.16\sqrt{10}$ を推定値とする。同様に、20 データポイントのピーク領域での標準偏差 $2.16\sqrt{20}$ は、サイコロを20回投げる実験を繰り返して推定する。

仮に、10 データポイントのピーク領域での3個の実験値と、20 データポイントのピーク領域での3個の実験値

を同等に扱ってみる。つまり、これらの6個の実験値から統計量を計算する。平均値はどちらも0であるから問題はないかもしれないが、6個の値からの標準偏差は、どちらのピーク領域での推定値なのか定義できない。計算値は必ず得られるが、どちらの真値（ $2\sqrt{10}$ 又は $2\sqrt{20}$ ）に対する推定なのかの解釈はできない。つまり、この計算方法は意味を成さない。

クロマトグラフィーでも、これと全く同じ議論が成り立つ。厳密に言えば、異なったピーク領域から計算した標準偏差は、ある母集団の標準偏差の推定値ではない。

4 結 論

クロマトグラムのピーク測定によく用いられる方法（インテグレーター法）の真度と精度を検討し、他の測定法（TOCO法）と比べた。インテグレーター法では、面積・高さ測定を行うためのピーク領域がベースラインノイズの変動に依存して変わるが、TOCO法では常に一定である。

インテグレーター法は、面積・高さ測定の両方で正のバイアスを与える。このため、検量線は正の切片を持つ。一

方, TOCO 法は, インテグレーター法に比べて, バイアスはかなり小さく, 検量線も原点を通る.

精度においては, これらの方法には, 大きな差はなかった.

どちらの方法でも, 検量線は直線であり, 定量分析の実践的観点からすれば, インテグレーター法も TOCO 法も問題はないと結論できる. なお, インテグレーター法の理論的問題点を指摘した.

文 献

- 1) S. R. Bakalyar, R. A. Henry: *J Chromatogr.*, **126**, 327 (1976).
- 2) H. Barth, E. Dallmeier, G. Courtois, H. E. Keller, B.

- L. Karger: *J. Chromatogr.*, **83**, 289 (1973).
- 3) J. P. Foley: *J Chromatogr.*, **384**, 301 (1987).
- 4) E. Grushka, I. Zamir: *Chem. Anal.*, **98**, 529 (1989).
- 5) Y. Hayashi, R. Matsuda: *Anal. Sci.*, **10**, 553 (1994).
- 6) A. N. Papas, M. F. Delaney: *Anal. Chem.*, **59**, 54A (1987).
- 7) J. B. Phillips: *Anal. Chem.*, **58**, 2091 (1986).
- 8) H. C. Smit, H. Steigstra: *ACS Symp. Ser.*, **361**, 126 (1988).
- 9) R. E. Synovec, E. S. Yeung: *Anal. Chem.*, **57**, 2162 (1985).
- 10) 林 譲, 松田りえ子: “HPLC 分析の精度”, (1999), (林純薬工業).
- 11) 四方田千佳子, 田頭洋子, 勝峰万里, 岩木和夫, 松田りえ子, 林 譲: 分析化学 (*Bunseki Kagaku*), **49**, 225 (2000).

要 旨

クロマトグラムのピーク測定によく用いられる方法 (インテグレーター法) と最近提案された方法 (TOCO 法) から得られる面積・高さの測定値の真度と精度を比較検討した. これらの方法の最も大きな違いは, インテグレーター法ではピーク領域がベースラインノイズの変動に依存して変わるが, TOCO 法では, ピーク領域 (幅) は常に一定であることにある.

クロマトグラフィーでよく観測されるベースラインノイズをコンピュータで合成し, ガウス型シグナルを重ね合わせた疑似クロマトグラムを作成し, インテグレーター法と TOCO 法で, 面積・高さを測定した. インテグレーター法は, 面積・高さ共に正のバイアスを与える. このため, 検量線は正の切片を持つ. 一方, TOCO 法は, インテグレーター法に比べてバイアスはかなり小さく, 検量線も原点を通る. 精度においては, パラメーターの設定によっては, インテグレーター法の精度は優れていることもあるが, 全体的には, これらの方法の精度には大きな差はない.

どちらの方法でも, 検量線は直線であり, 定量分析の観点からすれば, インテグレーター法も TOCO 法も問題はないと結論できる. インテグレーター法を用いて測定値から標準偏差を計算するときの理論的問題点を指摘する.