

Wprowadzenie

Poszukiwanie optymalnych warunków rozdzielania w chromatografii często bazuje na metodzie „prób i błędów”. W metodzie tej analityk integruje informacje z różnych źródeł, np. swoje doświadczenie, dane literaturowe oraz przeprowadzone już wcześniej eksperymenty wstępne, i na tej podstawie podejmuje dalsze kroki analityczne. Niestety, proces ten może być niedokładny, szczególnie jeżeli nie jest poparty rygorystycznym rozumowaniem. W pracy pokazano zastosowanie modelu hierarchicznego [1,2] oraz wnioskowania Bayesowskiego do podejmowania racjonalnych decyzji w kierunku uzyskania optymalnych warunków rozdzielania w układach RP-HPLC. Metoda ta pozwala przewidywać retencję analitów w oparciu o informację *a priori* (uzyskaną z analizy innych związków i opisaną przez stosowny model) oraz dowolną liczbę eksperymentów wstępnych dla nowych (nie analizowanych jeszcze) związków. Zastosowanie tej metodologii zostało zilustrowane na przykładzie 210 par analitów. Dla każdej pary szukano takiej zawartości modyfikatora organicznego w fazie ruchomej, która w warunkach izokratycznych pozwala uzyskać dużą rozdzielczość ($R > 1$) oraz współczynniki retencji związków w zakresie od 0.4 do 1 (4-12 min).

Metodyka

Wszystkie obliczenia zostały wykonane w programie Matlab (MathWorks, Natick, MA, USA)

Model

$$p(\log k | X, \theta, \sigma^2) = N(X\theta, \sigma^2)$$

$$X = [1 - \varphi / (1 + S_2 \varphi)] \text{ - gdzie } \varphi = [0.03:0.01:1]$$

$$\theta = [\log k_w S]^T \text{ - parametry specyficzne dla analitu}$$

$$t_R = t_0 + t_e + t_0 10^{\log k} \text{ - czas retencji}$$

$$L \text{ - długość kolumny}$$

$$W = 4 \sqrt{\frac{0.022 + \frac{(t_R - t_e)^2}{L} (0.00865 + \frac{0.136}{1 + \log k})}{1 + \log k}}$$

Rozkład *a priori*

$$p(\theta) = MVN \left(\begin{bmatrix} \theta_{\log k_w} + \theta_{\log k_w - \log P} \log P + \theta_{\log k_w - PSA} PSA \\ \theta_S + \theta_{S - \log P} \log P + \theta_{S - PSA} PSA \end{bmatrix}, \Omega \right)$$

Analyte Type	Theta (θ)	Omega (Ω)
Acid	$\begin{bmatrix} 0.43 + 0.915 \log P + 0.0144 PSA \\ 2.39 + 0.756 \log P + 0.0281 PSA \end{bmatrix}$	$\begin{pmatrix} 0.34106 & 0.24827 \\ 0.24827 & 0.9396 \end{pmatrix}$
Base	$\begin{bmatrix} -0.63 + 0.915 \log P + 0.0144 PSA \\ 1.559 + 0.756 \log P + 0.0281 PSA \end{bmatrix}$	$\begin{pmatrix} 0.34106 & 0.24827 \\ 0.24827 & 0.9396 \end{pmatrix}$
Neutral	$\begin{bmatrix} -0.63 + 0.915 \log P + 0.0144 PSA \\ 3.399 + 0.756 \log P + 0.0281 PSA \end{bmatrix}$	$\begin{pmatrix} 0.21716 & 0.24827 \\ 0.24827 & 0.43692 \end{pmatrix}$

Przewidywania *a priori*

$$p(\log \bar{k}^{(0)}) = \int p(\log k | X, \theta, \sigma^2) p(\theta) d\theta$$

$$E(\log \bar{k}^{(0)}) = X\theta$$

$$Var(\log \bar{k}^{(0)}) = X\Omega X^T + \sigma^2$$

Przewidywania *a posteriori*

$$p(\theta | \log k, X, \sigma^2) \propto p(\log k | \theta, X, \sigma^2) p(\theta)$$

$$\theta_n = E(\theta) = \Omega_n (\Omega_n^{-1} \theta + \sum_{j=1}^n \sigma_j^2 X_j^T \log k_j)$$

$$E(\log \bar{k}^{(n)}) = X\theta_n$$

$$p(\log k | \theta, X, \sigma^2) = \prod_{j=1}^n N(X_j \theta, \sigma_j^2)$$

$$\Omega_n = Cov(\theta) = \left(\Omega^{-1} + \sum_{j=1}^n \sigma_j^2 X_j^T X_j \right)^{-1}$$

$$Var(\log \bar{k}^{(n)}) = X\Omega_n X^T + \sigma^2$$

Prawdopodobieństwo sukcesu

$$Pr(\varphi^{(n)}) = Pr\{R > 1 \ \& \ 0.4 < \log \bar{k}_{Analiz}^{(n)} < 1 \ \& \ 0.4 < \log \bar{k}_{Analiz}^{(n)} < 1\}$$

$$R = 2 \frac{t_{R,1} - t_{R,2}}{W_1 + W_2}$$

Podejmowanie decyzji

Decyzja *a priori*: $Pr(\varphi^{(0)}) < 0.01$ (rozdzielenie niemożliwe)

Dla $Pr(\varphi^{(0)}) \geq 0.01$ przeprowadź eksperyment dla $\theta_{best}^{(0)}$

Stop jeżeli $\min(t_{R1}, t_{R2}) > 60$ lub $\max(t_{R1}, t_{R2}) > 120$ (rozdzielenie niemożliwe)

Jeżeli brak sukcesu

Decyzja *a posteriori*: $Pr(\varphi^{(1)}) < 0.05$ (rozdzielenie niemożliwe)

Dla $Pr(\varphi^{(1)}) \geq 0.95$ przeprowadź eksperyment dla $\theta_{best}^{(1)}$

itd

Symulacje:

5 000 par związków o następujących cechach:

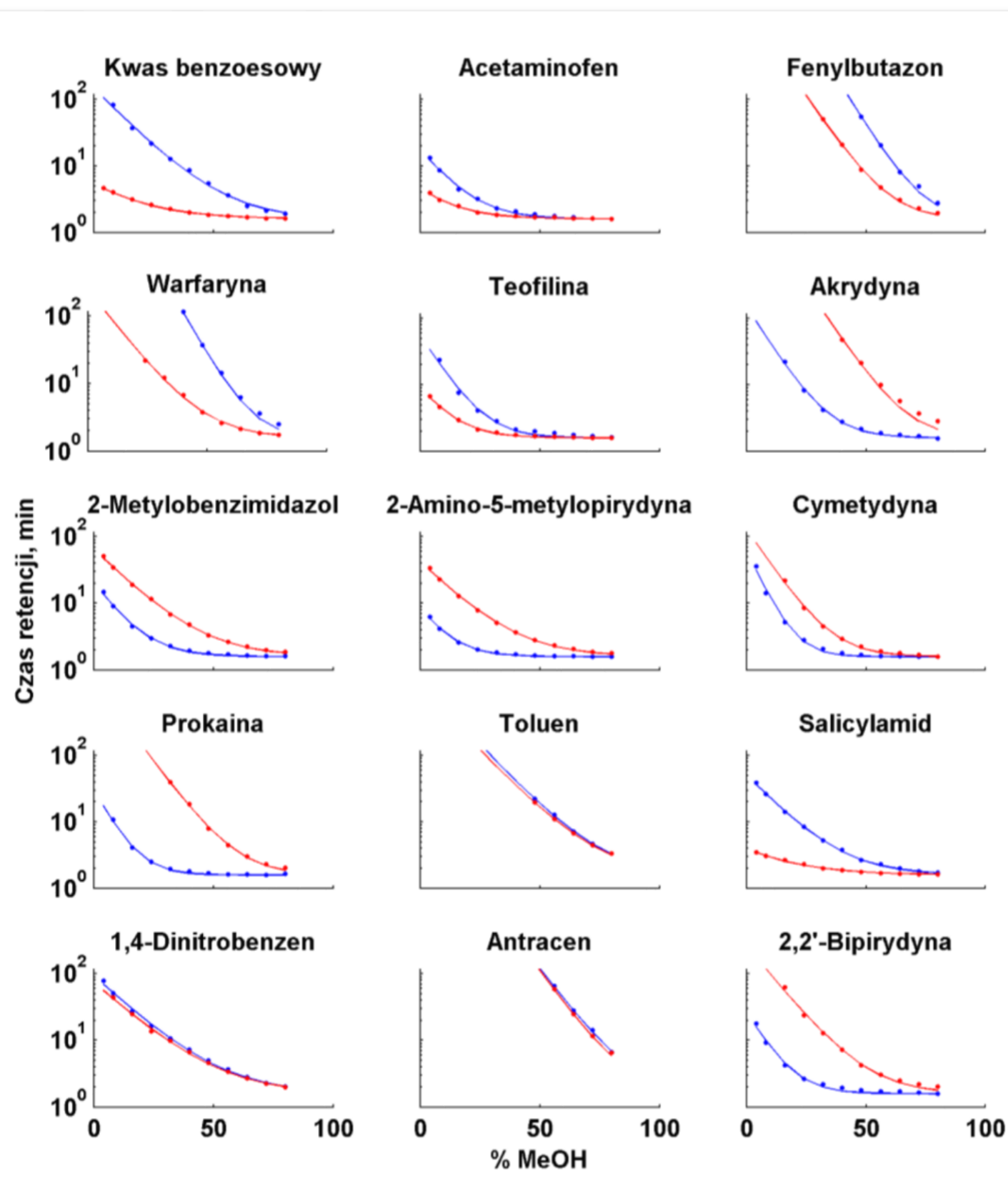
$$\log P \sim \text{uniform}(0.5, 4.5)$$

$$PSA \sim \text{uniform}(0, 100)$$

$$\text{Acid / Base / Neutral} \sim \text{multinomial}(1/3, 1/3, 1/3)$$

Pomiary:

Czasy retencji 15 związków oznaczono w warunkach izokratycznych w pH 2.50 (seria I) oraz pH = 10.5 (seria II). Do kontroli pH fazy ruchomej wykorzystano bufor uniwersalny: 0.008M kwas cytrynowy, glicynę, TRIS oraz 0.15M KCl. Eksperymenty przeprowadzono z wykorzystaniem aparatu Merck Hitachi oraz kolumny XTerra MS C18 5µm 4.6x150mm (Waters, USA). Szybkość przepływu fazy ruchomej wynosiła F = 1 ml/min, a temperatura T = 25°C.



Name	Kowariany			pH = 2.5		pH = 10.5	
	log P	pK _s	PSA	log _k (S)	S (%CV)	log _k (S)	S (%CV)
Kwas benzoesowy(A)	1.87	4.21	122.1	37	2.162 (3.26)	3.7	0.57 (2.66)
Acetaminofen (A)	0.339	9.38	151.2	53	1.199 (1.86)	4.5	0.48 (4.81)
Fenylbutazon (A)	3.16	4.5	308.4	41	4.732 (2.36)	6.8	1.44 (1.07)
Warfaryna(A)	3.417	4.82	308.3	64	4.68 (1.72)	7.7	2.39 (3.48)
Teofilina(A)	-0.175	8.81	180.2	69	1.728 (1.42)	5.8	0.68 (1.79)
Akrydyna(B)	3.40 ^a	5.45	179.2	13	2.168 (1.70)	5.7	3.79 (1.66)
2-Metylobenzimidazol (B)	1.575	6.19	132.2	29	0.844 (2.25)	5.4	1.66 (3.36)
2-Amino-5-metylopirydyna (B)	1.077	7.77	108.1	39	0.844 (1.80)	5.4	1.66 (1.94)
Cymetydyna (B)	0.261	5.23	680	252.3	114	1.772 (2.25)	2.14 (1.76)
Prokaina(B)	2.144 ^a	9.04	2.29	236.3	56	1.212 (1.82)	1.57 (3.47)
Toluen(N)	2.73	-	92.14	0	3.168 (1.01)	4.2	3.06 (1.94)
Salicylamid (A)	1.346	8.37	137.1	64	1.684 (0.89)	3.7	0.37 (5.83)
1,4-Dinitrobenzen (N)	1.367	-	188.1	92	1.956 (2.18)	3.5	1.85 (1.71)
Antracen (N)	4.55	-	178.2	0	3.646 (0.41)	5.9	4.74 (3.08)
2,2'-Bipirydyna(B)	1.28 ^a	4.33	165.2	26	1.285 (0.91)	5.9	2.42 (4.84)

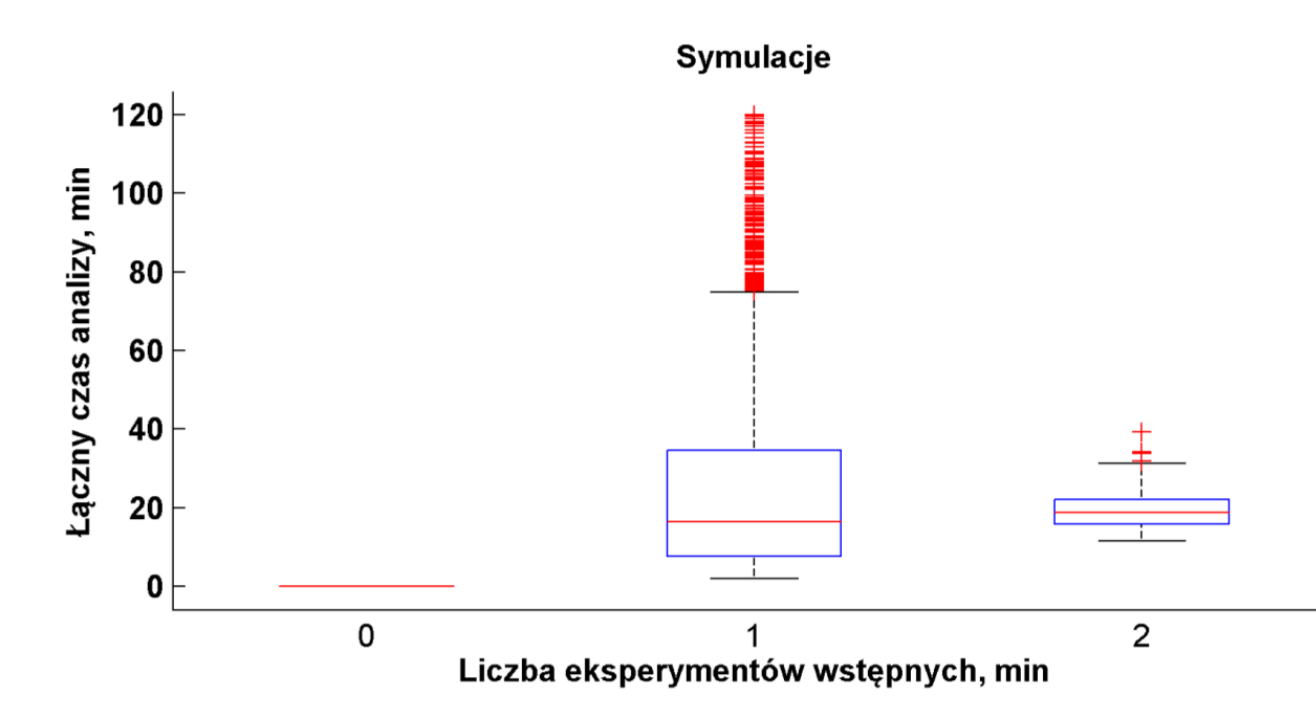
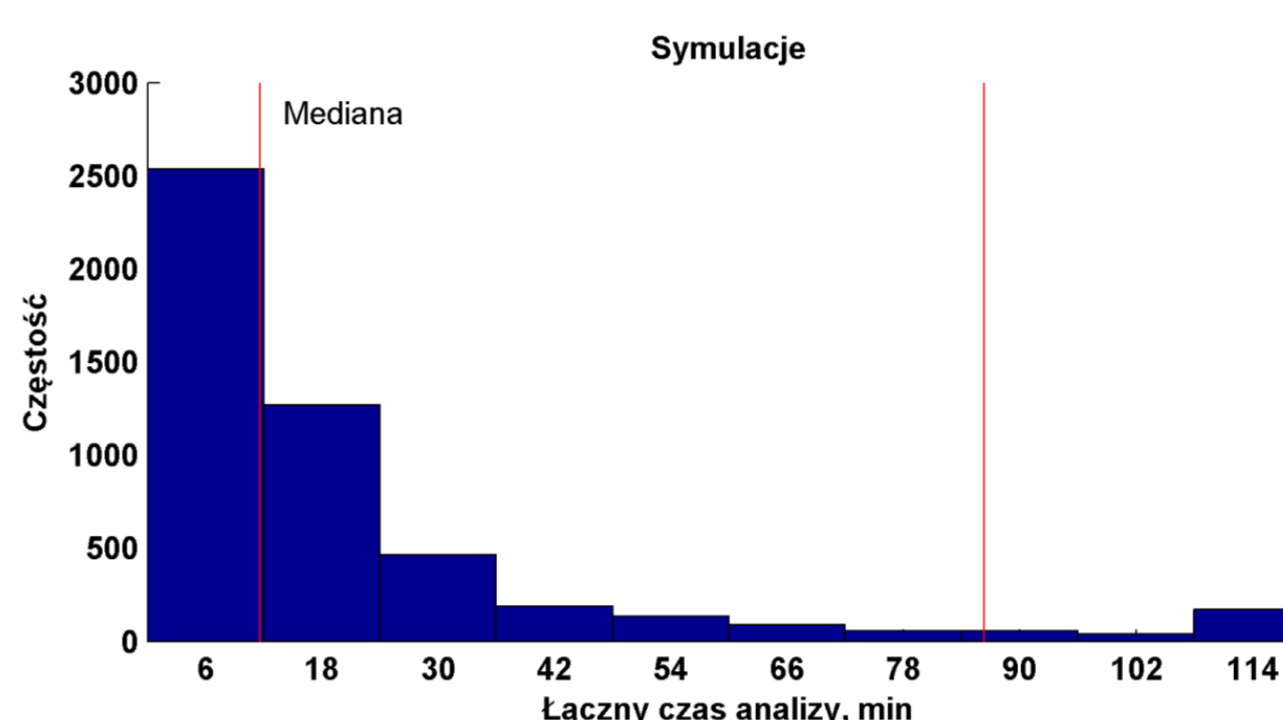
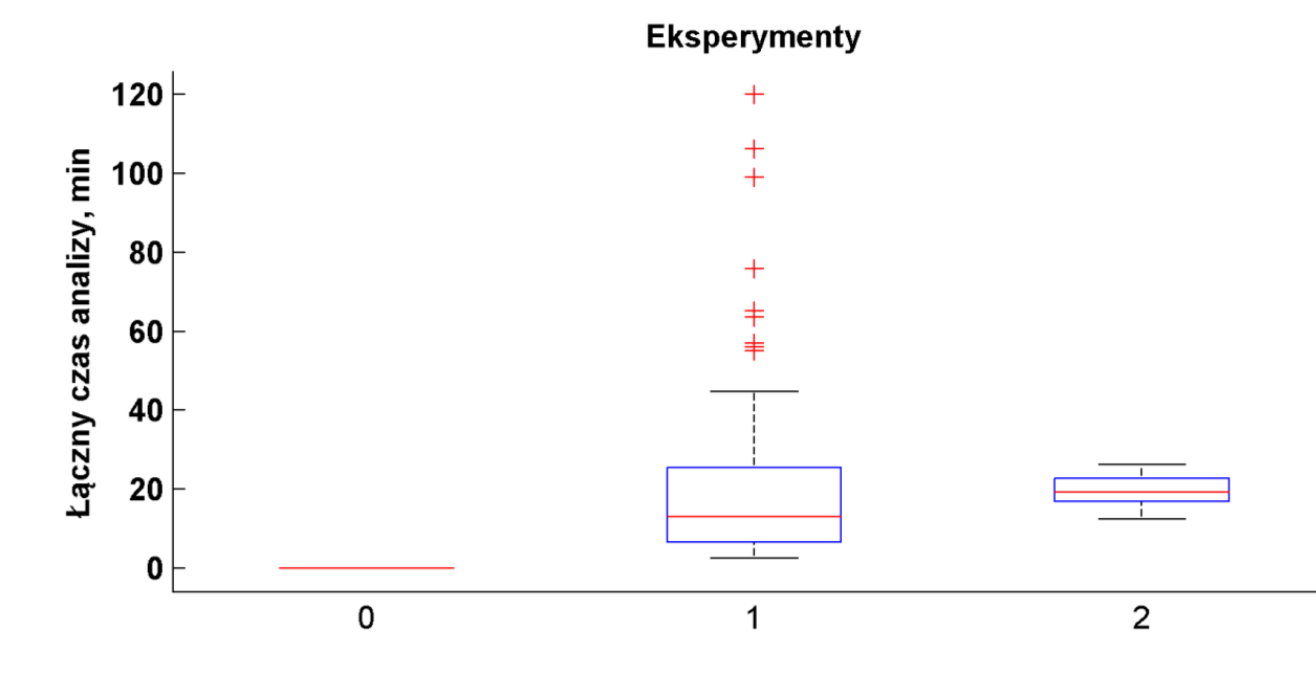
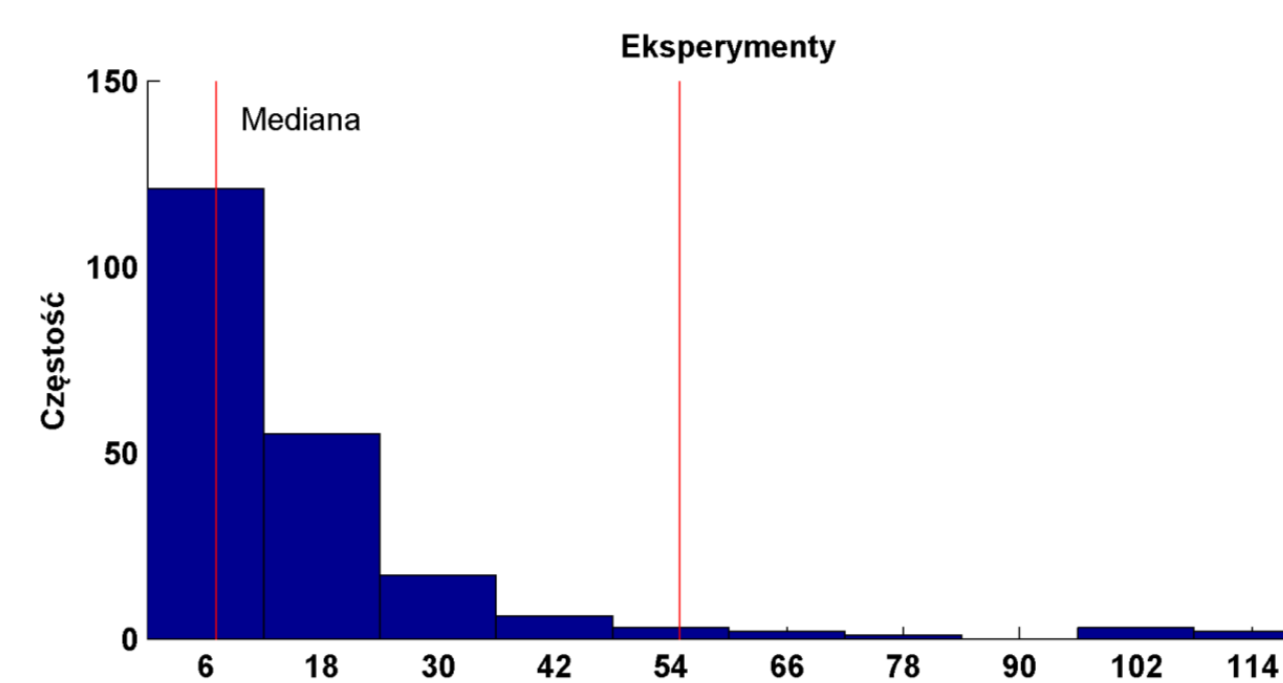
$$t_R = t_0 10^{\log k_w - \frac{S\varphi}{1+0.183\varphi}} + t_0 + t_e$$

t_R – czas retencji w pH = 2.50 lub pH = 10.50

t_e – czas poza kolumnowy (0.56 min)

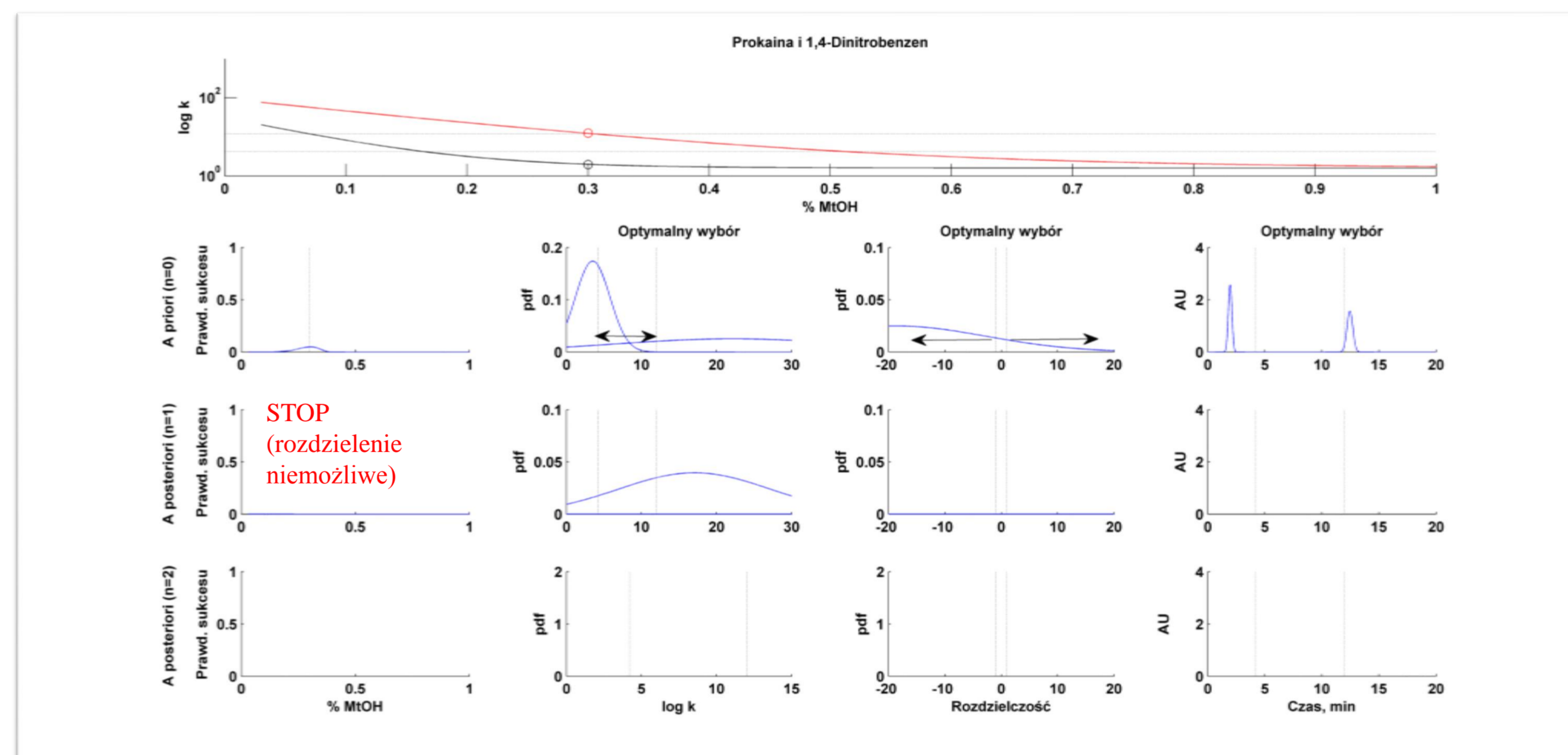
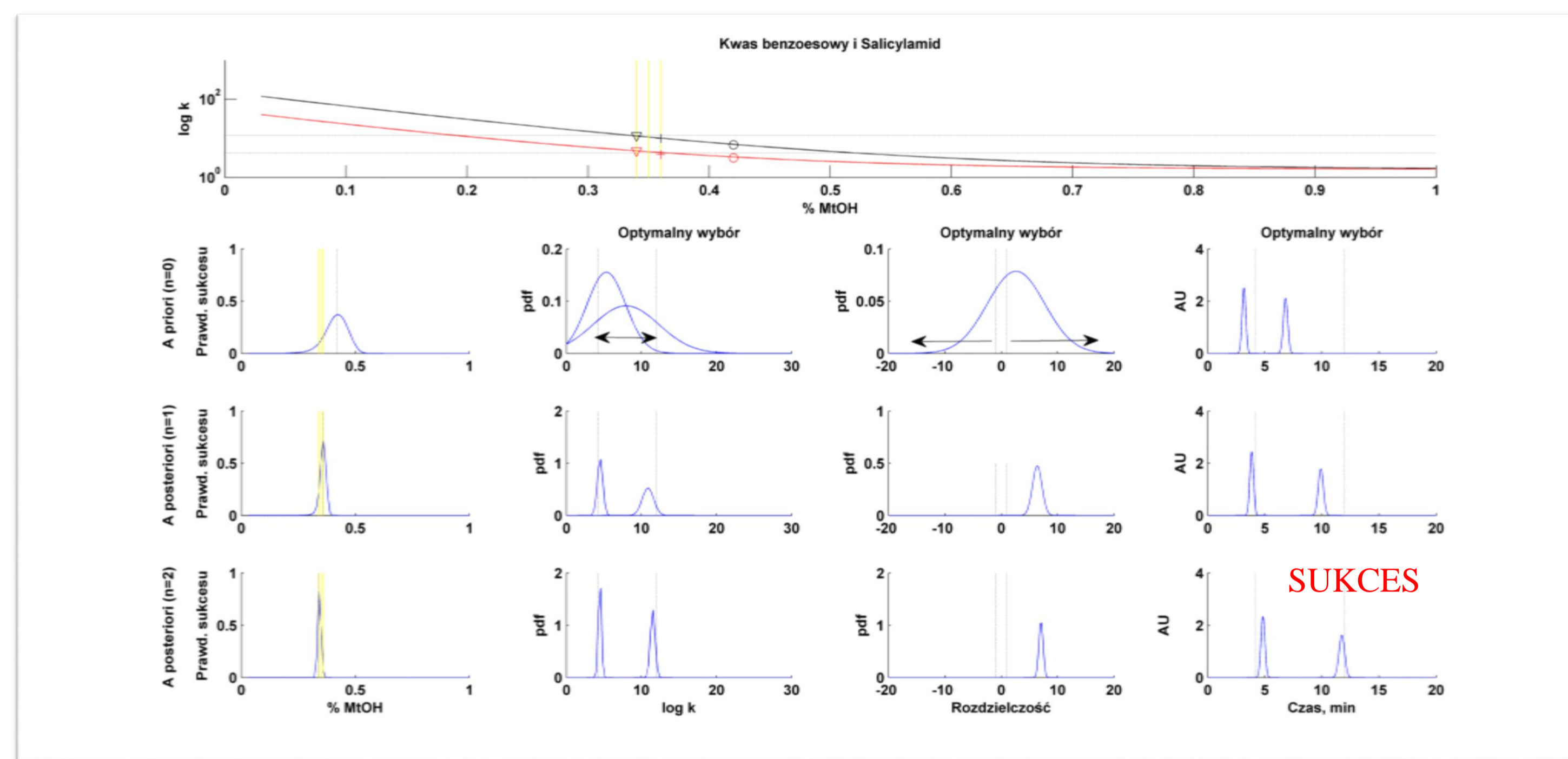
t_0 – czas martwy (1.04 min)

Wyniki



Macierz kontyngencji (eksperymenty)	Nieosiągnięta separacja (prawda)	Możliwa separacja (prawda)
Nieosiągnięta separacja (przewidywania)	152	2
Możliwa separacja (przewidywania)	2	54

Macierz kontyngencji (symulacje)	Nieosiągnięta separacja (prawda)	Możliwa separacja (prawda)
Nieosiągnięta separacja (przewidywania)	3556	48
Możliwa separacja (przewidywania)	71	1325



Wnioski

W przypadku testowanych związków żądaną separację uzyskano średnio w 7,1 min (0 – 55 min). Wymagało to 0, 1 i 2 eksperymentów wstępnych odpowiednio dla 30%, 60% i 10% par związków. Otrzymane wyniki sugerują, że zastosowanie technik wnioskowania Bayesowskiego daje możliwość szybkiego doboru warunków prowadzących do uzyskania żądanego rozdzielania (lub brak takiej możliwości) poprzez racjonalne wykorzystanie informacji dostępnej analitykowi.

Piśmiennictwo

- [1] Wiczling P, Kubik Ł, Kalisz R, Anal Chem, 2015, 87, 7241-9.
- [2] Wiczling P, Kalisz R, Anal Chem, 2016, 88, 997-1002.

Podziękowania