

15.1;07.2

## Нелинейно-регрессионный алгоритм обработки сигналов полупроводниковых химических сенсоров, обеспечивающий селективное детектирование примесей в искусственном воздухе

© В.В. Чистяков, С.А. Казаков, М.А. Гревцев, С.М. Соловьев

Физико-технический институт им. А.Ф. Иоффе РАН, Санкт-Петербург, Россия  
E-mail: v.chistyakov@mail.ioffe.ruПоступило в Редакцию 28 сентября 2020 г.  
В окончательной редакции 19 ноября 2020 г.  
Принято к публикации 1 декабря 2020 г.

Разработан новый способ обработки сигнала изменения электропроводности  $\Delta\sigma$  при температурной ( $T$ ) модуляции химического сенсора для селективного определения в воздухе следовых концентраций аммиака, ацетона, н-гексана, пропана, толуола и других примесей. Способ заключается в том, что в диапазоне прецизионно устанавливаемых концентраций  $C$  каждой из примесей  $Y$  сигнал  $\Delta\sigma$  как функция обратной температуры  $z = 10^3/T$  интерполируется при помощи нелинейной регрессии набором параметризованных функций  $F_i(z, A_i, b_i, c_i, \dots)$ ,  $i = 1-4$ , и строятся зависимости для главных (концентрационных) параметров  $A_{iY}(C)$ , определяющие так называемый портрет селективности  $Y$ . Вписываясь в него, аналогичные величины у детектируемой примеси  $X$  подтверждают ее тождественность с  $Y$ , а общая абсцисса всех точек пересечения линий уровня  $A_{iX}$  с кривыми  $A_{iY}(C)$  определяет численное значение и единицу измерения концентрации  $C_X$ .

**Ключевые слова:** химический сенсор, температурная модуляция, селективность, нелинейная регрессия, концентрация, углеводороды, аммиак, ацетон.

DOI: 10.21883/PJTF.2021.06.50751.18564

Ввиду высокой чувствительности электропроводности тонких пленок поликристаллических  $n$ -полупроводников ( $ZnO$ ,  $SnO_2$ ,  $SmS$  и др.) к хемосорбции из газовой среды на разветвленную поверхность как донорных ( $H_2$ ,  $CH_4$ ,  $CO$ ,  $C_2H_6$ ,  $NH_3$  и т.д.), так и акцепторных молекул ( $O_2$ ,  $O_3$ ) давно утвердилось представление о таких пленках как о перспективных материалах для чувствительных элементов химических сенсоров — приборов искусственного обоняния [1]. При этом проблема различения (селекции) различных веществ традиционно решается при помощи так называемой термо- или температурной модуляции, когда рабочая температура  $T$  [К] сенсора в среде изменяется в интервале  $(T_{low}, T_{up})$  монотонно либо циклическим образом [2–4].

В силу активационного характера хемосорбционных процессов температурная развертка должна давать различную кинетику прироста числа зонных носителей (электронов) и, следовательно, изменения во времени  $t$  проводимости  $\Delta\sigma(t)$  для частиц с разной энергетической структурой электронных оболочек, т.е. химией (см., например, [4], с. 1643).

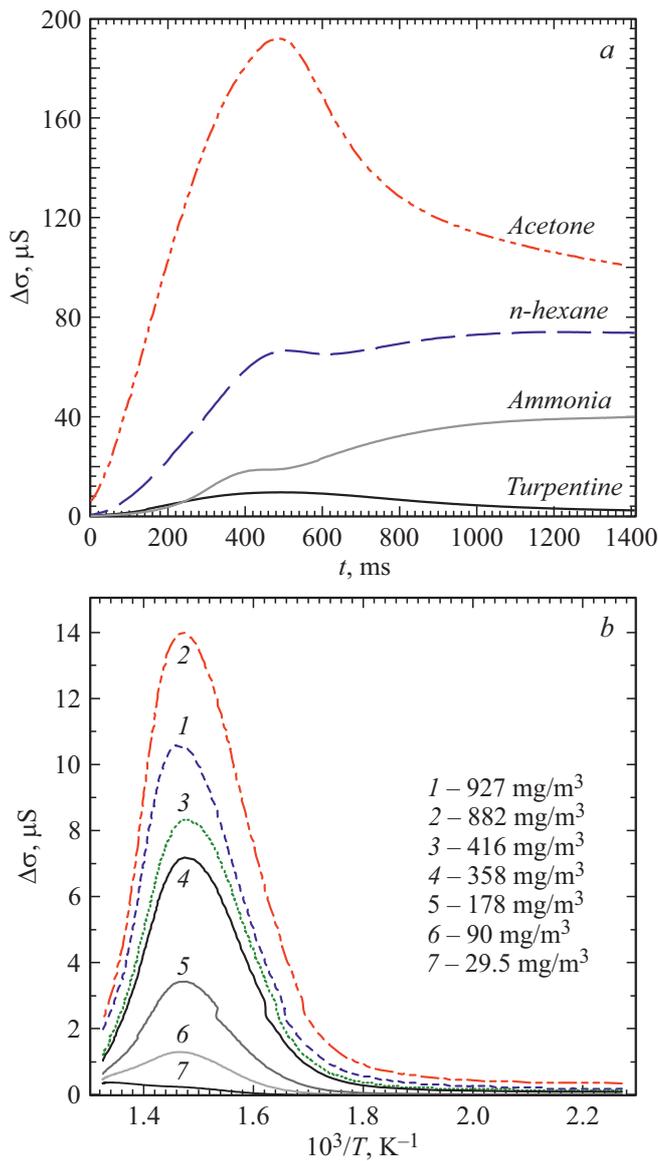
Математическим инструментом анализа здесь служит дискретное преобразование Фурье (быстрое преобразование Фурье) [2] временного ряда сигнала  $\Delta\sigma(t_i)$ ,  $i = 1-2^N$  при циклической термомодуляции с некоторым периодом  $\tau$ . При этом действительные части ( $R_k$ ,  $k = 0, 1, 2, \dots$ ) получаемых коэффициентов Фурье связывают с высотой приповерхностного потенциально-го барьера, а мнимые ( $I_k$ ,  $k = 1, 2, \dots$ ) — с величиной собственно адсорбции на поверхности [4]. Методика, однако, до сих пор не реализована в виде универсаль-

ного, надежного и простого компьютерного алгоритма распознавания примеси в анализируемой среде.

В настоящей работе предлагается иной подход, базирующийся на статистическом моделировании вышеозначенного сигнала. Подход дает контролируемые, объективные и в отличие от коэффициентов Фурье устойчивые к различным факторам интервальные, а не точечные оценки калибровочных характеристик. Опирающийся на аксиомы, выводы и формулы математической статистики, подход позволяет тем самым применение базирующихся на ней теории надежности, теории принятия решений и вообще искусственного интеллекта. Поэтому он может быть реализован в виде ясного алгоритма с использованием коммерческого программного обеспечения и затем воплощен в реальных сенсорных системах с передачей сигнала через Bluetooth или Wi-Fi на сервер автоматической обработки, анализа и принятия решений.

Метод строится на соображениях, что сигналы рассматриваемого сенсора от разных адсорбатов (рис. 1) при температурной развертке („пиле“) несут в себе различия, выражающиеся как явно — в положении особых точек (максимумов, минимумов, перегибов) (рис. 1,  $a$ ), так и неявно, кумулятивным образом, — во всех остальных уровнях временного/температурного ряда  $\Delta\sigma(t_k/t_k^0)$ ,  $k = 1, \dots, N = 400 \dots 800$ .

Прежде всего для всех адсорбатов группы выбирают один и тот же график нагрева. Аргументом сигнала  $\Delta\sigma(z)$  формально служит величина  $z = 10^3/T$  [К<sup>-1</sup>] (рис. 1,  $b$ ).



**Рис. 1.** Отклики прироста  $\Delta\sigma$  электропроводности сенсора. *a* — для одинаковой ( $\sim 400 \text{ mg}/\text{m}^3$ ) концентрации разных веществ относительно времени экспозиции  $t$ ; *b* — для различных концентраций паров сольвента относительно величины  $z = 10^3/T$ .

Определяемые исходя из отклика величины должны обладать свойством робастности, т.е. устойчивости к 1) замене на эквивалентный набор данных, 2) одиночным выбросам (своеобразным дискретным „ $\delta$ -функциям“, искажающим весь спектр Фурье) и 3) шумам сигнала. В качестве таковых выбирались значимо оцениваемые параметры формальной нелинейной по параметрам и объясняющей переменной регрессии сигнала  $\Delta\sigma(z)$  с использованием ряда дискриминирующих функций  $F_n(z)$ :

$$F_1(z) = \frac{0.01A_1}{z^4 + b_1z^3 + c_1z^2}, \quad (1)$$

$$F_2(z) = \frac{0.01A_2}{z^3 + b_2z^2 + c_2z}, \quad (2)$$

$$F_3(z) = \frac{0.01A_3}{z^4 + b_3z^3 + c_3z^2 + d_3z}, \quad (3)$$

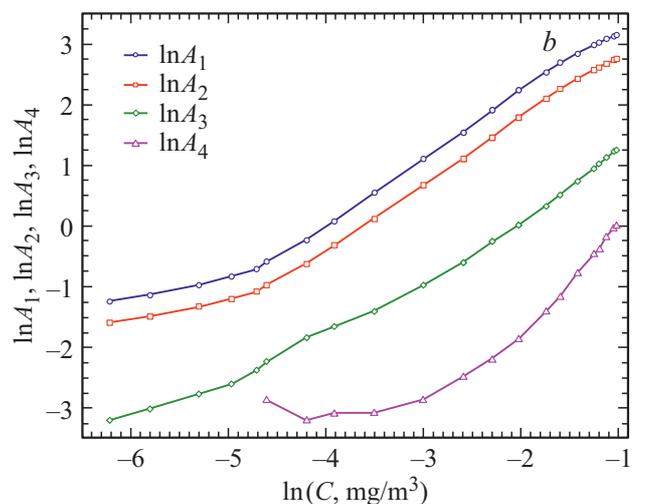
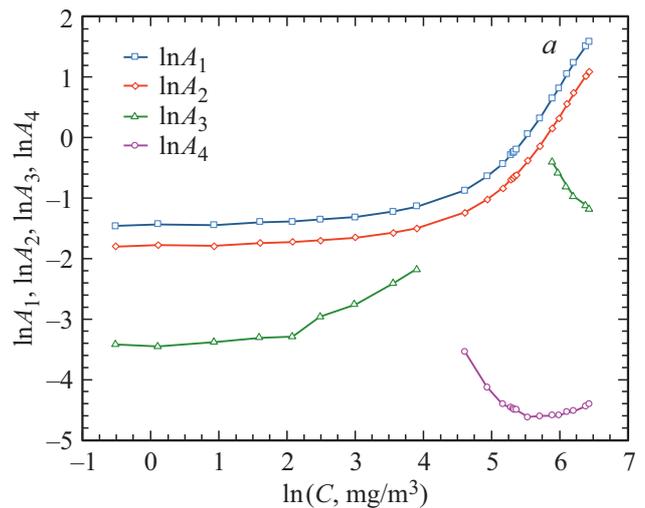
$$F_4(z) = \frac{0.0001A_4}{z^4 + b_4z^3 + c_4z^2 + d_4z + h_4}, \quad (4)$$

$$F_5(z) = \frac{0.01A_5}{b_5z^4 + c_5z^3 + d_5z^2 + h_5z + 1} \quad (5)$$

(множители 0.01/0.0001 использованы для более точной оценки малых значений главных параметров  $A_i$  — центров соответствующих доверительных интервалов (ДИ) на доверительном уровне/вероятности  $P = 0.95$ ).

В идеале все величины  $A_i$  должны расти с концентрацией  $C$ . Но у одних монотонность утрачивается в силу немонотонности размаха отклика  $\Delta\sigma$  на пары жидкостей, у других — из-за неизбежных флуктуаций получаемых оценок или иных причин.

Необходимо выполнение следующих требований к зависимостям  $A_i(C)$ : 1) положительно определенность;



**Рис. 2.** Портреты селективности (gauge portraits) различных веществ в искусственном воздухе для сенсора на основе  $\text{SnO}_2$ . *a* — аммиак, *b* — пропан.

**Таблица 1.** Результаты оценки по формуле (1) параметров сигнала  $\Delta\sigma/\mu\text{S}$  SnO<sub>2</sub>-сенсора на концентрацию  $C = 951 \text{ mg/m}^3$  толуола (доверительный уровень  $P = 0.95$  (95%))

Параметр	$adj-R^2 = 0.73335699$				
	Оценка	Стандартная ошибка	Критерий Стьюдента $t$	Нижняя граница	Верхняя граница
$A_1$	7.05398	0.253780	27.80	6.55546	7.55251
$b_1$	-3.08519	0.001573	-1961.59	-3.08827	-3.08210
$c_1$	2.38329	0.002524	944.40	2.37833	2.38824

**Таблица 2.** Результаты оценки по формуле (1) параметров сигнала  $\Delta\sigma/\mu\text{S}$  SnO<sub>2</sub>-сенсора на концентрацию  $C = 45 \text{ mg/m}^3$  толуола (доверительный уровень  $P = 0.95$  (95%))

Параметр	$adj-R^2 = 0.99579415$				
	Оценка	Стандартная ошибка	Критерий Стьюдента $t$	Нижняя граница	Верхняя граница
$A_1$	0.10740	0.002137	50.27	0.10320	0.11160
$b$	-2.93652	0.001519	-1933.04	-2.93950	-2.93354
$c_1$	2.16033	0.002168	996.48	2.15607	2.16458

2) визуальной непрерывности; 3) среди них обязательно должны быть монотонные в рабочем диапазоне концентраций или обладающие ярко выраженным растущим трендом и при этом 4) в координатах  $\ln A_i - \ln C$  имеющие широкий участок крутого наклона.

В силу робастности значимых регрессионных оценок вообще фактически отпадают строгие требования на интервал температурной развертки  $\Delta T$ , тем более что границы диапазона регрессора  $z$  слабо чувствительны к разумным изменениям уровней температурной „пилы“: на 10–20 К для нижнего  $T_{low}$  и на 20–30 К для верхнего  $T_{up} = T_{low} + \Delta T$ .

По структуре дискриминирующих функций (1)–(5) главные параметры  $A_i(C)$  должны быстро меняться с концентрацией  $C$  в противоположность медленно меняющимся  $b_i(C)$ ,  $c_i(C)$ ,  $d_i(C)$ ,  $h_i(C)$ , играющим вспомогательную идентифицирующую роль.

Кроме того, такую же роль могут играть также исправленный коэффициент детерминации  $adj-R^2$ , показывающий в какой мере изменение сигнала  $\Delta\sigma$  обусловлено изменением регрессора  $z = 10^3/T$ , и величины критерия Стьюдента оценок  $t(A_i)$ ,  $t(b_i)$ ,  $t(c_i)$ , ..., т.е. отношение сигнал/шум для всех оцениваемых параметров.

Основная идея метода заключается в том, что для каждой примеси из некоторой группы (аммиак, н-гексан, этанол, ацетон, сольвент, толуол и пр.) вышеозначенным образом оцениваются и строятся калибровочные кривые  $A_i(C)$ ,  $i = 1-4$  или 5, тем самым для каждого вещества формируется уникальный портрет селективности. Такой же набор  $\{A_{iX}, i = 1-4 \text{ или } 5\}$  также путем регрессии по формулам (1)–(5) находится и для неизвестной концентрации неизвестной же примеси  $X$ . Если он (набор) вписывается в портрет селективности некоторого вещества  $Y$ , т.е. все горизонтальные линии  $A_{iX}$  пересекаются со всеми соответствующими кривыми  $A_{iY}(C)$

на портрете  $Y$  в точках с равными (в идеале) абсциссами  $C^*$ , то детектируемая примесь и есть  $Y$ , а  $C^*$  — ее концентрация в единицах портрета  $Y$ .

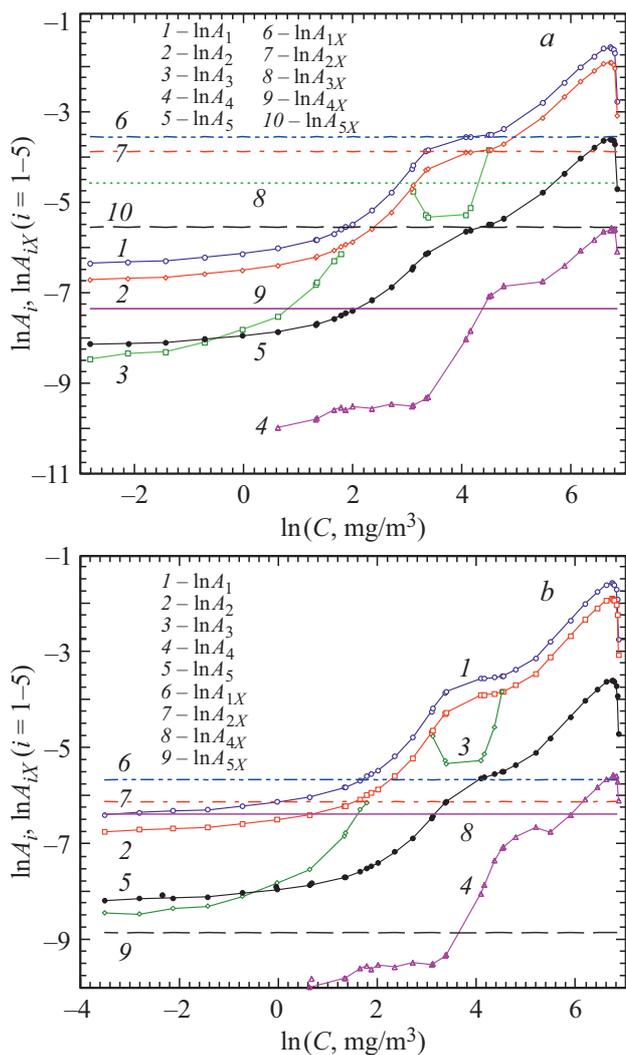
Параметры оценивались опцией Nonlinear Estimation продукта StatSoft Statistica Release 12 (Trial, см. [5]), высокая оценка эффективности которого для обработки данных физических экспериментов дана в [6]. При этом минимизировалась стандартная сумма квадратов отклонений (loss function), а в качестве вычислительных алгоритмов служили методы Гаусса–Ньютона и Левенберга–Марквардта [5].

Начальные значения оцениваемых параметров, как правило, полагались нулевыми, но их можно получить из линейной по коэффициентам множественной регрессии для обратной величины сигнала  $1/\Delta\sigma$  [MΩ] в соответствии с модельными зависимостями  $\frac{1}{F_i(z)}$  (1)–(5) ( $\frac{1}{F_i(z)} = \alpha z^4 + \beta z^3 + \gamma z^2$ ) с последующим пересчетом найденных коэффициентов.

В работе анализировались массивы данных по изменению сигнала электропроводности при температурной развертке от  $\sim 430$  до  $\sim 750$  К примерно за время  $\Delta\tau \sim 1.4$  s, полученные для тонких пленок poly-SnO<sub>2</sub>(CuO), изготавливаемых по золь-гель-технологии [1].

Функции (1)–(5) описывали экспериментальные данные при малых концентрациях с высоким коэффициентом  $adj-R^2 = 0.99 \dots 0.999$ , но с ростом  $C$  величины  $adj-R^2$ , как правило, монотонно плавно убывали до 0.73–0.80. При этом критерии Стьюдента оставались на высоком уровне: для главных параметров  $t(A_i)$  порядка 15–150, для остальных коэффициентов  $t(b_i)$ ,  $t(c_i)$ ,  $t(d_i)$ ,  $t(h_i)$  — 200–2000 (табл. 1, 2).

Параметр считался оцененным значимо на уровне  $P = 0.95$  ( $\alpha = 1 - P = 0.05$ ), только если на таком



**Рис. 3.** Идентификация сольвент–не сольвент при помощи портрета селективности. *a* — подтверждение в концентрации 72.2–79.8  $\text{mg/m}^3$  (все линии уровня  $\ln A_{iX}$  главных параметров для примеси  $X$  пересекаются с соответствующими кривыми  $\ln A_i(C)$  для сольвента в точках с одинаковыми абсциссами); *b* — неподтверждение (абсциссы точек разные, для линии  $\ln A_{5X}$  пересечения нет вообще, величина  $\ln A_{3X}$  значимо не оценена).

уровне была значима вся их совокупность. Вследствие же незначимости на кривых  $A_3(C)$  и  $A_4(C)$  имели место пробелы. Параметры  $A_1(C)$ ,  $A_2(C)$  и  $A_5(C)$  значимо оценивались у всех веществ и при всех  $C$  [ $\text{mg/m}^3$ , vol.%], и именно их линии служили калибровочными кривыми.

Были получены различающиеся по виду наборы регулярных зависимостей  $\ln A_i(\ln C)$  (портреты селективности) для следующих веществ в искусственном воздухе: аммиака, ацетона, н-гексана, пропана, сольвента и др. (рис. 2, 3); затем они использовались для безошибочной и точной идентификации газов из данного списка.

Пример такой идентификации паров неизвестного газа  $X$  (со следующим ДИ  $A_i$ :  $A_{1X \text{ low}} = 2.81870 < A_{1X} < A_{1X \text{ up}} = 2.90253, \dots, A_{5X \text{ low}} = 0.36973 < A_{5X} < A_{5X \text{ up}}$

$= 0.39618$ ), оказавшегося сольвентом в концентрации  $\sim 78 \text{ mg/m}^3$ , схематично показан на рис. 3, *a*. Здесь горизонтальные линии центров ДИ  $A_{iX}$  для газа  $X$  пересекаются с кривыми портрета  $Y$  (сольвент) в точках с визуально одинаковыми абсциссами  $C^*$ . В спорных случаях необходимо будет оперировать уже интервалами концентраций  $C$ , границы которых для каждой точки пересечения ( $i$ ) суть графические релевантные решения „уравнений“  $A_{iY \text{ up}}(C) = A_{iX \text{ low}}$  и  $A_{iY \text{ low}}(C) = A_{iX \text{ up}}$ . В том случае, если пересечение всех таких интервалов не есть пустое множество, детектируемая примесь соответствует портрету.

На рис. 3, *b* представлен обратный случай отрицательной идентификации неизвестного газа  $X'$  (толуола). Оба примера свидетельствуют о простоте и эффективности нелинейно-регрессионного метода. И даже немонотонность откликов сенсора по концентрации (рис. 3, *a*) не является здесь серьезной помехой для идентификации и измерения.

Таким образом, разработанный метод позволяет уверенно идентифицировать донорные микропримеси в искусственном воздухе. Его возможности и ресурсы далеко не исчерпаны как в плане оптимизации по параметрам развертки, так и в плане применения искусственного интеллекта [7].

Метод уже сейчас может быть обобщен на детектирование двух примесей ( $X_1, X_2$ ), например пропана и паров воды. На предварительно построенных 3D-портретах  $A_i(C_1, C_2)$  линии уровня параметров  $A_{iX_1, X_2}$ , будучи спроецированы на плоскость концентраций ( $C_{X_1}, C_{X_2}$ ), пересекутся там в одной точке, которая и определит состав двойной примеси.

### Финансирование работы

Исследование выполнено в рамках гранта Российского фонда фундаментальных исследований № 18-03-00660.

### Конфликт интересов

Авторы заявляют, что у них нет конфликта интересов.

### Список литературы

- [1] K.S. Shalini Devi, A. Anantharamakrishnan, U. Maheswari Krishnan, J. Yakhmi, in *Smart sensors for environmental and medical applications* (IEEE, Inc., 2020), p. 103–127. DOI: 10.1002/9781119587422
- [2] S. Nakata, S. Akakabe, M. Nakasuji, K. Yoshikawa, *Anal. Chem.*, **68**, 2067 (1996).
- [3] A. Heilig, N. Bârsan, U. Weimar, M. Scheizer-Berberich, J.W. Gardner, W. Göpel, *Sensors Actuators B*, **43**, 46 (1997).
- [4] S. Nakata, T. Hashimoto, H. Okunishi, *Analyst*, **127**, 1642 (2002). DOI: 10.1039/b208295k
- [5] <https://www.statsoft.de>
- [6] В.В. Чистяков, *Физическое образование в вузах*, **21** (1), 120 (2015).
- [7] P. Gwiźdź, A. Brudniak, K. Zakrewska, *Metrol. Meas. Syst.*, **XXII** (1), 3 (2015). DOI: 10.1515/mms-2015-0008