

ЛИНЕЕН МОДЕЛ НА ИЗХОДНАТА МОЩНОСТ НА CuBr ЛАЗЕР С ПОМОЩТА НА CART

Десислава Войникова

Факултет по математика и информатика, Пловдивски университет, България
voynikova@uni-plovdiv.bg

LINEAR MODEL OF THE OUTPUT POWER OF CuBr LASER USING CART

Desislava Voynikova

Faculty of Mathematics and Informatics, University of Plovdiv, Bulgaria
voynikova@uni-plovdiv.bg

***Abstract.** This study examines copper bromide vapor laser with wavelengths 510.6 nm and 578.2 nm. It has various practical applications. It is determined as the most effective laser in the visible range and is of great scientific interest. Based on all available experimental data for this type of laser, which has been developed and patented by the Laboratory of metal vapor lasers at the Institute of Solid State Physics, BAS, there was made a study with the nonparametric technique CART (Classification and Regression Trees). The binary recursive method used for classification is a powerful tool for complex structure. There is obtained regression model based on CART, which describes over 95% of data. The built regression tree of the output laser power can be used to classify existing and new observations.*

***Key words:** CuBr laser, CART, output power, binary tree splitting, v-fold cross-validation*

1. Въведение

В техническите системи съществуват голям брой величини – геометрични, физични, описващи свойствата на използваните материали и други, които до голяма степен определят различни изходни характеристики на системата. Лазерните устройства функционират в зависимост от над 50 параметри със сложни връзки между тях, описващи протичащите процеси най-често с голям брой диференциални и друг тип уравнения. Решавайки ги, получаваме числови данни за решенията. Такъв подход се нарича структурен подход. В същото време при наличие на данни от експерименти може да се подходи в обратната посока, което се нарича феноменологичен подход. При него целта е на базата на данните да извлечем основните зависимости (уравнения, връзки) за изходните характеристики. Това най-често се постига с помощта на статистически методи, генетични и еволюционни алгоритми и други.

В настоящата работа се използва вторият подход. За наличните данни за лазер с пãри на меден бромид се прилага методът CART (Classification and Regression Trees), с който наблюденията се разделят на групи във вид на двоично дърво според изходната характеристика – изходна лазерна мощност. Ще отбележим, че за този лазер са получени както параметрични, така и непараметрични модели. В (Gocheva-Пиева, Пиев 2011) са построени и изследвани различни линейни и нелинейни модели с помощта на

метода на многомерна регресия с главните елементи, за случайни извадки за изходната мощност и за лазерната ефективност. Непараметрични модели бяха получени с метода на многомерните регресионни сплайни в (Iliev, Voynikova, Gocheva-Ilieva 2012), които описват над 98% от експерименталните данни с относителна точност съизмерима с точността на измерванията.

Целта на това изследване е да се проведе анализ и прогнозиране на експерименти на базата на полученото решение от CART за изходната лазерна мощност.

Резултатите са получени с помощта на софтуерния пакет CART.

2. Описание на данните

За предмета на настоящата работа са използвани данни от 387 проведени експеримента с CuBr лазери. Основната зависима променлива е средната изходна лазерна мощност P_{out} (W). Като предиктори са използвани следните 6 входни лазерни променливи: D (mm) – вътрешен диаметър на лазерната тръба; P_{in} (kW) – входна електрическа мощност; $P_L = P_{in}/L$ (kW/cm) – електрическа мощност на единица дължина с 50% загуба; P_{rf} (kHz) – честота на повторение на електрическия импулс; p_{H_2} (Тorr) – налягане на водорода; C (nF) – еквивалентен капацитет на кондензаторната батерия. Описателните статистики на извадката са показани в таблица 1. От наблюденията за асиметрия и ексцес може да се заключи, че данните нямат нормално разпределение. В такъв случай се препоръчват непараметрични регресионни техники, за които няма такова изискване.

Таблица 1. Статистически характеристики за набор от $n = 387$ измервания.

Променлива	Минимум	Максимум	Средна с-ст	Асиметрия	Ексцес
D	15.00	58.00	46.59 ± 10.072	-0.809 ± 0.12	1.451 ± 0.25
Pin	30.00	200.00	106.59 ± 70.70	0.478 ± 0.12	-1.670 ± 0.25
P_L	5.00	16.67	10.92 ± 2.51	0.467 ± 0.12	0.183 ± 0.25
P_{rf}	0.00	0.80	0.36 ± 0.25	-0.416 ± 0.12	-1.430 ± 0.25
p_{H_2}	8.00	250.00	22.56 ± 24.17	6.389 ± 0.12	46.454 ± 0.25
C	0.33	4.00	1.33 ± 0.61	2.313 ± 0.12	6.233 ± 0.25
P_{out}	0.25	120.00	34.024 ± 35.57	0.808 ± 0.12	-0.862 ± 0.25

При предишно проучване (Iliev, Voynikova, Gocheva-Ilieva 2012) на данните за лазер с пари на CuBr с метод на многомерните адаптивни регресионни сплайни (MARS) е установено, че тези независими променливи са от съществена важност при построяване на модел за изходната лазерна мощност P_{out} .

Използвани са експериментални данни за лазери с пари на меден бромид, направени през последните десетилетия в Лабораторията по метални пари, Институт по физика на твърдото тяло „Георги Наджаков”, Българска академия на науките - София.

Данните са публикувани в (Sabotinov 2006) и цитираната там литература за този лазер.

3. Кратко описание на метода CART

Алгоритъмът CART, както се вижда от наименованието, решава задачата за класификация и регресия. Разработен е през 1974-1984 г. от Leo Breiman, Jerry Friedman, Charles Stone и Richard Olshen. (Breiman, Friedman, Olshen, Stone 1984)

CART е непараметрична техника за дърво на решенията, която построява класификационни или регресионни дървета, в зависимост от това дали зависимата променлива е съответно категорийна или числова.

Алгоритъмът е предназначен за построяване на двоично (бинарно) дърво на решенията, като на всяка стъпка от разрастването (за всеки негов възел) се използва правило, което дели съвкупността от наблюдения на две подсъвкупности. Целта е намиране на дърво, което извършва добро разделяне на данните, по избран критерий. Всеки клон на дървото завършва в терминален възел, като всяко наблюдение попада в точно един краен възел, и всеки такъв е еднозначно определен от набора от правила.

По-конкретно, цел в подхода на регресионното дърво е разделяне на данните в относително хомогенни (с ниско стандартно отклонение) крайни възли и получаване на средна наблюдавана стойност при всеки възел като прогнозна стойност. Изграждането на дървото започва с първото двоично разделяне на най-важната променлива, наречен възел-родител. Впоследствие тези възли се разделят многократно до достигане на краен възел. За всеки възел, променливата, която осигурява най-добро разделяне с определен критерий при минимална грешка е включена в дървото.

Тъй като дърветата могат да бъдат чувствителни към случайна грешка в данните, при изграждане на регресионни дървета обикновено се използва процесът на валидиране. Техниката за валидиране, използвана от CART – крос-валидация, дава възможност за изграждане на много надеждни модели, които са по-добри от стандартните регресионни модели. За целите на това проучване е използвана стандартна крос-валидация. Данните са разделени по случаен принцип на 10 равни непресичащи се подчасти, всяка съдържаща 10% от съвкупността. Дървото е построено с помощта на 9/10 от данните (обучителна извадка), а останалите 1/10 (тестова извадка) е използвана за прогнозиране и за определяне на нивото на грешка. Процесът по изграждане на дървото се повтаря 10 пъти и средната грешка от 10-те серии се разглежда като обща оценка. Тази процедура осигурява точно оценяване на зависимата променлива и дава възможност построеното дърво да се използва за друг набор от данни.

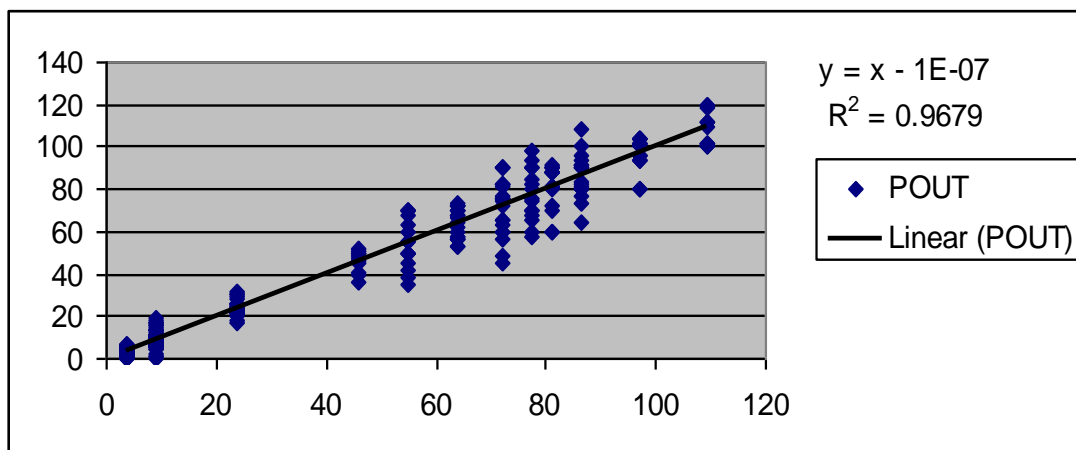
4. Построяване на линеен CART модел за изходната мощност

Построен е CART модел за установяване зависимостта на лазерната изходна мощност P_{out} от 6-те основни входни независими лазерни променливи. Зададено е ограничение за минимум 20 наблюдения във възел-родител и минимум 10 за краен възел. Това ще гарантира, че крайните възли няма да са твърде малки.

Целта е да се разкрие групирането, което води до по-висока изходна лазерна мощност. Като CART може да се използва както за анализ, така и за предсказване на стойностите на зависимата променлива.

За оценяване на качеството на разделяне е използван методът на средноквадратична грешка с най-малките квадрати.

От избраното за оптимално дърво, показано на фигура 1 се вижда, че първи, втори и трети крайни възли включват експериментите с най-ниска изходна лазерна мощност. Възли 4, 5, 6 и 8 групират наблюдения със средни резултати за P_{out} , а възли 7, 9-13 включват експериментите с висока изходна лазерна мощност. Конкретните условия за разделяне на всеки възел се виждат от таблица 2.



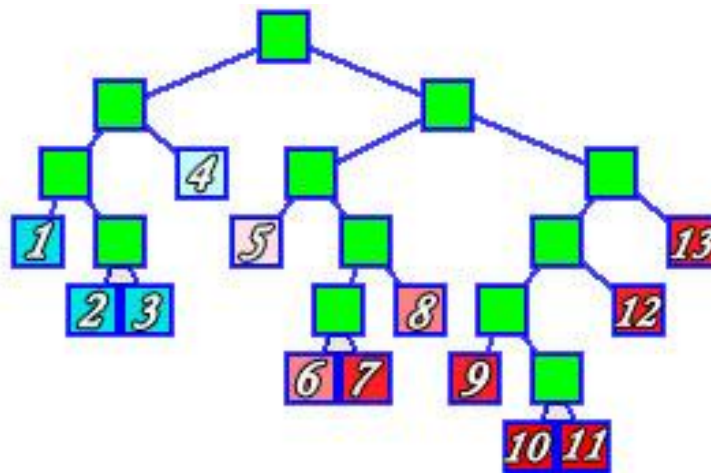
Фигура 1. Сравнение на реалните данни с предсказанияте от полученото CART дърво на решенията.

Таблица 2. Основна информация за дърво с 13 крайни възела.

Крайни възли	Минимум	Максимум	Средна стойност	Брой	Класификационни условия за възела
1	0.4	6.2	3.827	99	$P_{in} \leq 1.95$ $P_{in} \leq 1.35$ $PH2 \leq 0.05$
2	1	6.7	3.730	15	$P_{in} \leq 1.95$ $P_{in} \leq 1.35$ $PH2 > 0.05$ $D \leq 27.50$
3	0.25	19	9.071	94	$P_{in} \leq 1.95$ $P_{in} \leq 1.35$ $PH2 > 0.05$ $D > 27.50$
4	16.8	32	23.937	35	$P_{in} \leq 1.95$ $P_{in} > 1.35$
5	36	51.8	46.000	11	$P_{in} > 1.95$ $P_{in} \leq 3.15$ $P_{in} \leq 2.45$
6	53	73	63.910	20	$P_{in} > 1.95$ $P_{in} \leq 3.15$ $P_{in} > 2.45$ $C \leq 1.45$ $P_{in} \leq 2.85$
7	60	92	81.188	16	$P_{in} > 1.95$ $P_{in} \leq 3.15$ $P_{in} > 2.45$ $C \leq 1.45$ $P_{in} > 2.85$
8	34	70	54.733	15	$P_{in} > 1.95$ $P_{in} \leq 3.15$ $P_{in} > 2.45$ $C > 1.45$
9	64	108	86.375	16	$P_{in} > 1.95$ $P_{in} > 3.15$ $C \leq 1.75$ $PRF \leq 19.25$ $PRF \leq 15.50$
10	80	104	97.000	16	$P_{in} > 1.95$ $P_{in} > 3.15$ $C \leq 1.75$ $PRF \leq 19.25$ $PRF > 15.50$ $P_{in} \leq 4.25$
11	100	120	109.286	14	$P_{in} > 1.95$ $P_{in} > 3.15$ $C \leq 1.75$ $PRF \leq 19.25$ $PRF > 15.50$ $P_{in} > 4.25$
12	58	98	77.400	15	$P_{in} > 1.95$ $P_{in} > 3.15$ $C \leq 1.75$ $PRF > 19.25$
13	45	90	72.048	21	$P_{in} > 1.95$ $P_{in} > 3.15$ $C > 1.75$

Моделът с 13 крайни възела има качество на приближение $R^2 = 0.96791$, т.е. той описва добре около 97% от всички данни. Относителната грешка е 0.038 или се постига почти перфектно съвпадение. Като от построеното дърво е видно, че най-висока изходна лазерна мощност има при експерименти с ограничение на входните променливи $P_{in} > 4.25$ kW, $C \leq 1.75$ nF и 15.5 kHz $< P_{rf} \leq 19.25$ kHz. На фигура 2 е

направено сравнение между експерименталните и предсказаните стойности по модела за изходната лазерна мощност.



Фигура 2. Дърво на решенията с 13 крайни възли.

Посредством регресионното дърво, съгласно условията за принадлежност на наблюденията към локализиращия значим краен възел (11-ти) с най-висока очаквана средна стойност, може да се извърши прогнозиране, че при бъдещи експерименти, наблюденията класифицирани в тази група ще имат стойности за P_{out} в интервала от 100 до 120 W. Като предсказаната средна стойност на изходната мощност е 109.286 W.

5. Заключение

В настоящото изследване на зависимостта на изходната мощност на лазери с пари на меден бромид от шест входни променливи, са класифицирани наблюденията по групи. Избрани са правила, въз основа на стойностите на променливите, така че да се получи най-доброто разделяне на наблюденията спрямо зависимата променлива. Използвайки регресионните дървета, CART може да бъде приложен както за анализ, така и за предсказване на нови наблюдения, като осъществяването на прогнози се извършва посредством вътрешни проверки, базирани на това дали променливата се намира под или над определена стойност за дадена променлива.

Благодарности

Това изследване е финансово подпомогнато от проект НИ11-ФМИ 004 към НПД на Пловдивски университет „Паисий Хилендарски”.

Литература

Gocheva-Ilieva, S., Iliev, I., Statistical models of characteristics of metal vapor lasers. // Nova Science Publ., Inc., New York, 2011, pp. 194.

Iliev, I., Voynikova, D., Gocheva-Ilieva, S., Simulation of the output power of copper bromide lasers by the MARS method. // Quantum Electronics, 2012, 42, N 4, pp. 298-303.

Sabotinov, N., Metal vapor lasers, in: M. Endo, R.F. Walter (Eds.). // Gas Lasers, CRC Press, Boca Raton, 2006, pp. 449-494.

Breiman, L., Friedman, J., Olshen, R., Stone, C., Classification and Regression Trees, Wadsworth International, Canada, 1984.

CART® Classification and Regression Trees. 2012. <<http://www.salford-systems.com/en/products/cart>> 15.10.2012