

ПРАКТИЧЕСКО ПРИЛОЖЕНИЕ НА СТАТИСТИЧЕСКИ ТЕХНИКИ В МАТЕРИАЛОЗНАНИЕТО

Мартин Иванов¹, Николай Тончев²
martinivanov@abv.bg, tontchev@vtu.bg,

¹*Нов Български Университет*

²*Висше Транспортно Училище „Тодор Каблешков“ - София
БЪЛГАРИЯ*

Ключови думи: *Моделиране на свойства, статистически пресмятания, магнезиеви сплави.*

Резюме: *В изследването се прилагат стандартни статистически процедури за определяне на експертна оценка за влиянието на елементите от състава на сплавта върху контролираните механични свойства – якост на опън и относително удължение. Изведени са невронни модели, описващи механичните характеристики от количеството на алуминия, мангана, никела, желязото и силиция с пряко значение към експлоатационните свойства на изделието. Чрез приложения подход е възможно да се определи състав, осигуряващ относително оптимални стойности на изследваните индикатори на качеството.*

1. УВОД

Решаването на съвременните технически задачи се свързва с отстраняване на дадено техническо или физично противоречие на факторите, описващи проблема. В областта на материалознанието - това е противоречието между якост – R_m [MPa] и относително удължение- A [%]. Множеството от техни алтернативи се получава чрез съответно легиране или термично обработване. Така се формира задача, в която входните параметри са: състава, определен от количеството на конкретния елемент и параметрите на режима на обработване. За решаване на подобни задачи се търсят подходи, улесняващи лицето вземащо решение ЛВР /Decision Maker/ в анализите, които той/тя правят. Различно получените решения са свързани с различен определен ефект на постигнати ползи. За да се приложат статистически техники, за анализ на избран индикатор на качеството, е необходимо да се изведат модели за изследваните величини при проведен експеримент. Позовавайки се на богатата информация от компаниите [1] - [4], на различни бази от данни при изследване на зависимостите от тях е възможно да се направят ценни предложения подобряващи параметрите на създаваните изделия и технологии.

Изследването е част от серия развиващи се подходи и методики [5], които в етапа на генериране на решението не използват знания, натрупани в областта на металургията. То е логическо продължение на експеримент, проведен във връзка с националната програма за високи технологии, изследвания и развитие – Програма – 863 на Китай през 2009 – проект AA03Z525 на фонда за науката и технологиите на град Далиан и вътрешен проект J21DW003/2009 .

Цитираното изследване [5] е методология, явяваща се като начин за проектиране на сплави при предварително зададена база от данни. Именно за това тук се поставя възможността да се илюстрират използването на съвременни изчислителни средства за решаване на специфични инженерни задачи в областта на обработка на магнезиевите сплави.

2. МЕТОДИКА ЗА ПРОВЕЖДАНЕ НА ИЗСЛЕДВАНЕТО

Методиката за провеждане на изследването се състои от следните етапи:

- Предварителен статистически анализ на данните от изследването с визуализация на зависимостите между наблюдаваните величини. Това включва определяне на базовите (описателни) статистически характеристики, наличието на корелация на параметрите на експерименталните изследвания.
- Апроксимация на зависимостите между участващите в химичния състав на сплавта химични елементи и механичните ѝ характеристики с помощта на невронни модели.

Предмет на анализа е съвкупност от експериментални данни за състава и свойствата на магнезиевите сплави. Данните съдържат 41 експериментални наблюдения. Цел на анализа е установяване на основните статистически характеристики, описващи данните от експеримента, както и съставяне на статистически модели, апроксимиращи връзките между химическия състав на сплавта и физико-механичните параметри якост на опън R_m и относително удължение A .

Резултатите от предварителния статистически анализ на наличните експериментални данни, включващ изчисляването на основни дескриптивни статистики, са показани в табл. 1. Те включват: средна стойност, минимална и максимална стойност, стандартно отклонение.

Таблица 1. Основни дескриптивни статистики за наблюдаваните променливи.

Променлива	Средна стойност	Минимум	Максимум	Стандартно отклонение
Al	5,2159	0,0000	10,700	3,24134
Mn	0,3585	0,0800	1,5000	0,25968
Zn	1,2139	0,0000	6,5000	1,78290
Cu	0,2510	0,0000	2,7000	0,62465
Fe	0,0046	0,0000	0,0500	0,01304
Ni	0,0059	0,0000	0,0300	0,00928
Si	0,2588	0,0000	1,0000	0,33983
R_m	228,707	150,00	295,00	37,5887
A	7,2439	2,0000	15,000	3,50557

3. СТАТИСТИЧЕСКИ АНАЛИЗ НА ДАННИТЕ И ВИЗУАЛИЗАЦИЯ

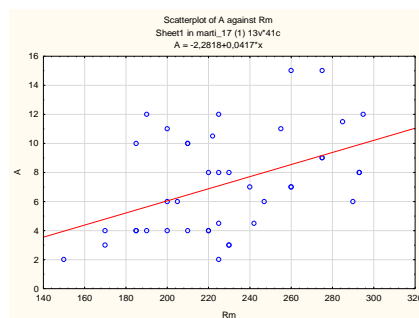
Най-съществената задача на този етап е да се установи възможност за намиране на евентуална връзка на независимите параметри и зависимите характеристики в експерименталното изследване. Статистическият анализ позволява да бъдат определени некорелираните входни параметри от експеримента, които могат да се използват за построяването на регресионен модел – процентното съдържание на елементите в състава на магнезиевата сплавта. Ограниченият брой на наблюденията не дава възможност да се направят изводи относно статистическото разпределение на наблюдаваните променливи.

Стойностите на статистически определените корелационни коефициенти между наблюдаваните променливи (две по две) са показани в табл. 2. Като цяло те имат ниски стойности, повечето от които са статистически незначими (при p-level = 0.05). Изключение правят няколко от стойностите, които в табл.2. са отбелязани с удебелен и наклонен (bold+italic) шрифт.

Таблица 2. Коефициенти на корелация между променливите.

Променливи	Al	Mn	Zn	Cu	Fe	Ni	Si	Rm	A
Al	1,000000	-0,620126	-0,471530	-0,446552	0,367573	0,180073	-0,013168	-0,021270	-0,324544
Mn	-0,620126	1,000000	0,289916	0,229913	-0,047708	-0,216358	-0,100364	0,030075	-0,021431
Zn	-0,471530	0,289916	1,000000	0,863515	-0,111306	-0,006764	-0,255629	0,157414	-0,034116
Cu	-0,446552	0,229913	0,863515	1,000000	-0,043246	0,064104	-0,117705	0,022081	-0,159779
Fe	0,367573	-0,047708	-0,111306	-0,043246	1,000000	0,201250	0,062923	-0,270441	-0,224916
Ni	0,180073	-0,216358	-0,006764	0,064104	0,201250	1,000000	0,285475	-0,075156	-0,087665
Si	-0,013168	-0,100364	-0,255629	-0,117705	0,062923	0,285475	1,000000	-0,324950	-0,229324
Rm	-0,021270	0,030075	0,157414	0,022081	-0,270441	-0,075156	-0,324950	1,000000	0,446598
A	-0,324544	-0,021431	-0,034116	-0,159779	-0,224916	-0,087665	-0,229324	0,446598	1,000000

Интерес представлява коефициентът на корелация между променливите Rm и A поради възможността за включването им в общ многокритериален модел. Този корелационен коефициент е положителен и статистически значим със стойност равна на 0,446598. На фиг. 1. е показана диаграма на взаимната корелация между стойностите на параметрите Rm и A.



Фиг. 1. Диаграма на корелираните стойности на променливите Rm и A.

4. МОДЕЛИРАНЕ НА СТАТИСТИЧЕСКИТЕ ЗАВИСИМОСТИ С ЛИНЕЕН РЕГРЕСИОНЕН МОДЕЛ

Опитът да се състави линеен многомерен регресионен модел, обвързващ предикторите с резултативните променливи Rm и A не даде задоволителен резултат и значими статистически оценки за коефициентите в модела и за общата значимост на регресионния модел - най-вероятно поради ограничения брой на наблюденията, както и поради отсъствие на ясно изразени линейни връзки между предикторите и резултативните променливи в модела. Получените стойности на коефициентите и в двата модела в по-голямата си част, както и за общата оценка от ANOVA-анализа на моделите се оказват статистически незначими (при p-level = 0.05). Построяването на линеен регресионен модел не е удачно решение за апроксимиране на връзките между променливите в експеримента.

Използване на невронни модели за апроксимация на връзките между предикторите и резултативните променливи Rm и A.

Възникналите проблеми с построяването на статистическите модели, апроксимиращи влиянието на предикторите върху стойностите на резултативните стойности на Rm и A породиха необходимостта да се търси моделирането на изследваните връзки с други средства. За решаването на конкретната задача бе избран

подходът с използването на невронни модели от типа „многослоен перцептрон“ (ММР). Този подход не представлява новост, но се оказва твърде ефективен в случаите когато класическите статистически методи не дават резултат. Той дава възможност за съставяне на апроксимационни модели в случаите когато връзките между наблюдаваните величини са значително по-сложни, а понякога и неяви. По принцип невронните модели не се отнасят към статистическите техники, а се разглеждат като самостоятелни резултати от машинното обучение (machine learning). Въпреки това много автори отбелязват изчислителното сходство на тези модели с методите на статистическия анализ.

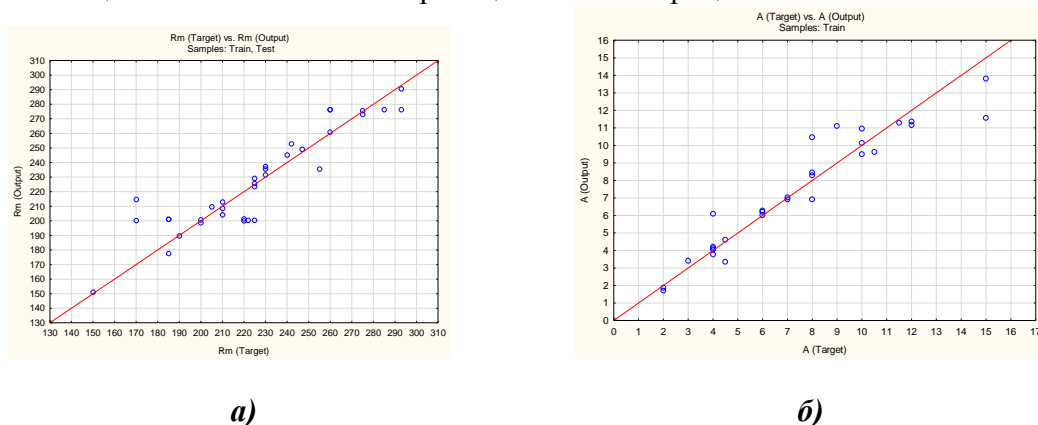
Съставянето на невронни модели, описващи адекватно връзките между предикторите и физико-механичните параметри Rm и A бе изпълнено с модула Automatic Neural Network на популярния статистически пакет Statistica 12. Проведен е експеримент с над 2000 модела за всяка от търсените апроксимации, като подборът, е извършен според стойността на корелационните коефициенти между наблюдаваните и изчислените резултативни стойности (съответно за обучаващото, тестовото и валидиращото множество). Определените като най-добри невронни модели, както и показателите за качеството им са показани в табл.3.

Таблица 3. Показатели за качеството на избраните невронни модели за апроксимация на Rm и A.

Апроксимиран параметър	Тип мрежа	Обучение	Тестване	Валидиране	Алгоритъм	Грешка	Активационна функция скрит слой	Активационна функция изход
Rm	MLP 7-6-1	0,947569	0,8861244	0,80383108	BFGS 93	SOS	Logistic	Identity
A	MLP 7-10-1	0,952151	0,9030694	0,90535722	BFGS 70	SOS	Logistic	Tanh

В анализа на резултатите от числения експеримент по моделирането на изследваните връзки са включени следните стъпки и оценки:

- Съставяне на двумерни диаграми за разпределение на наблюдаваните (по хоризонталната ос) и моделираните (по вертикалната ос) стойности. Визуалната интерпретация на диаграмата показва значителна линейна корелация между тези величини и за двата модела (якостта на опън и относителното удължение), съответстващи на стойностите на корелационните коефициенти от табл.3.



Фиг. 2. Диаграми за разпределение на наблюдаваните (по хоризонталната ос) и моделираните (по вертикалната ос) стойности, а) – Rm, б) – A.

Бяха съставени на хистограми на разпределението на остатъците (разликите между наблюдавани и моделирани стойности). Статистическите параметри на тези разпределения са съответно : за Rm – средна стойност -1,03232 и стандартно отклонение 14,21009; за A – средна стойност -0,14856 и стандартно отклонение 1,36316.

Опитите, статистическите разпределения на остатъците, да бъдат апроксимирани с нормално разпределение не дадоха задоволителен резултат. Прилагането на Chi-square критерия показва стойности под стандартното ниво на значимост p-level = 0.05, което е статистически признак за значимо различие между наблюдаваното и теоретичното разпределения

Анализът на резултатите от апроксимацията с посочените невронни модели бе допълнен и с анализ на чувствителност. Последният показва относителната значимост на предикторите във формирането на стойностите на резултатните променливи. За съставените модели тези резултати са показани в табл. 4 . Признак за значимо влияние на предиктора е ако стойността му е по-голяма от 1.

Таблица 4. Анализ на чувствителност за моделите, апроксимиращи Rm и A.

	Модел	Zn	Si	Cu	Al	Fe	Mn	Ni
Rm	MLP 7-6-1	92,063	37,877	21,068	11,916	6,6961	6,467061	4,927017
A	MLP 7-10-1	12,926	9,3068	5,54604	16,482	6,70571	2,58310	12,1877

Въз основа на данните от анализа на чувствителност може да се направи изводът, че най-силно влияние върху параметъра Rm оказва елементът цинк, а за A най-силно влияещият елемент е алуминий.

5. ЗАКЛЮЧЕНИЕ.

Приложена е статистическа експертна оценка за влиянието на елементите от състава на сплавта върху предварително избрани индикатори на качеството с цел подобряване механичните свойства на изделията. Изведени са невронни модели, описващи механичните характеристики на количеството на легиращите елементи в магнезиевите сплави , които са от значение за експлоатационните свойства на изделието. Чрез приложението подход е възможно да се определи състав, осигуряващ относително най-добри значения на стойностите на селектираните механични показатели.

ЛИТЕРАТУРА:

- [1] Friedrich H.E., Mordike B.L. Magnesium Technology: Metallurgy, Design, Data, Applications, Springer-Verlag Berlin Heidelberg, 2006. 647 p.
- [2] Kainer K.U. (Ed.) Magnesium Alloys and Technologies Wiley-VCH, 2003. 285.
- [3] Monteiro W.A. (Ed.) New Features on Magnesium Alloys, InTech, 2012. 174 p.
- [4] Czerwinski F. (Ed.) Magnesium Alloys: InTAe, 2014, 176 p.
- [5] Tontchev N. Materials Science, Effective solutions and Technological variants, 2014/3/3, LAMBERT Academic Publishing.

PRACTICAL APPLICATION OF STATISTICAL TECHNIQUES IN MATERIAS SCIENCE

Martin Ivanov, Nikolay Tonchev
tontchev@gmail.com, martinivanov@abv.bg

¹ *New Bulgarian University – Sofia*
² *Todor Kableshkov University of Transport, Sofia,*
BULGARIA

Key words: *Materials Science, Modelling of properties, ANN , magnesium alloys.*

Abstract: *Standard statistical procedures for determining expert judgment for the influence of alloy composition elements on the controlled mechanical properties - tensile strength and relative elongation are used in the study. Neural models describing the mechanical characteristics of the aluminum, manganese, nickel, iron and silicon, of direct importance to the product's performance properties, are derived. Through the applied approach, it is possible to determine a composition providing relatively optimal values of the quality indicators investigated.*