



**ЮГОЗАПАДЕН УНИВЕРСИТЕТ
“НЕОФИТ РИЛСКИ” - БЛАГОЕВГРАД
ПРИРОДО-МАТЕМАТИЧЕСКИ ФАКУЛТЕТ
КАТЕДРА “ИНФОРМАТИКА”**

Радослав Стефанов Мавревски

**МОДЕЛИРАНЕ И ОПТИМИЗАЦИЯ ПРИ
МЕДИКО-БИОЛОГИЧНИ ИЗСЛЕДВАНИЯ**

АВТОРЕФЕРАТ НА ДИСЕРТАЦИЯ

*ЗА ПРИСЪЖДАНЕ НА ОБРАЗОВАТЕЛНА И НАУЧНА СТЕПЕН
„ДОКТОР”*

Област на висше образование 4. Природни науки, математика и информатика

Професионално направление 4.6. Информатика и компютърни науки

Докторска програма “Информатика”

Научни ръководители: проф. д-р Петър Миланов
доц. д-р Георги Георгиев

БЛАГОЕВГРАД, 2015 г.

Дисертационният труд е обсъден и насочен за публична защита на заседание на катедра „Информатика“ при ПМФ на ЮЗУ „Неофит Рилски“, състояло се на 21.04.2015 г.

Дисертацията съдържа 140 страници и се състои от използвани съкращения, увод, три глави, изводи, списък на цитираната литература и приложения. Включени са 30 фигури и 14 таблици. Цитираната литература обхваща 192 заглавия, от които 14 на кирилица и 178 на латиница.

Защитата на дисертационния труд ще се състои на 17.07.2015 г. от 11:00 часа в зала 1412 на УК-1 на ЮЗУ „Неофит Рилски“, Благоевград.

Материалите по защитата са на разположение на интересувашите се в кабинет № 1426 в УК-1.

Автор: Радослав Стефанов Мавревски
Тема: МОДЕЛИРАНЕ И ОПТИМИЗАЦИЯ ПРИ МЕДИКО-БИОЛОГИЧНИ ИЗСЛЕДВАНИЯ

СЪДЪРЖАНИЕ

	<i>стр.</i>
ОБЩА ХАРАКТЕРИСТИКА НА ДИСЕРТАЦИОННИЯ ТРУД	04
1. Актуалност и мотивировка на темата	04
2. Цели и задачи на дисертацията	06
3. Предмет на изследването	07
4. Методи на изследване	07
5. Научна новост	08
6. Аprobация	09
7. Публикации	09
СЪДЪРЖАНИЕ НА ДИСЕРТАЦИОННИЯ ТРУД	10
ПЪРВА ГЛАВА: ПРОБЛЕМАТИКА И ОБОСНОВКА НА ИЗСЛЕДВАНИЯТА	10
ВТОРА ГЛАВА: МЕТОДИЧНИ ПОДХОДИ	20
ТРЕТА ГЛАВА: РЕЗУЛТАТИ И ДИСКУСИЯ	26
ИЗВОДИ	43
СПИСЪК С ПУБЛИКАЦИИ ПО ДИСЕРТАЦИЯТА	44
УЧАСТИЕ В ИНТЕНЗИВНИ КУРСОВЕ И УЧАСТИЕ В ПРОЕКТИ	45
ЛИТЕРАТУРА	46
ANNOTATION IN ENGLISH	49

ОБЩА ХАРАКТЕРИСТИКА НА ДИСЕРТАЦИОННИЯ ТРУД

1. Актуалност и мотивировка на темата

Медико-биологичните изследвания са свързани с проектирането на биологични експерименти, събиране, обобщаване и анализ на данни от тези експерименти, както и тълкуването на резултатите и изводите от тях. Статистическите методи, математическото и компютърното моделиране са от критично значение за медико-биологичните изследвания. Обединяването на целта на моделирането и експериментални изследвания при медико-биологичните науки е свързано с изясняване на основните биологични процеси, в резултат на дадени наблюдавани явления. Важно предизвикателство за математическото моделиране включва определяне на подходящ модел от даден клас, който да бъде използван за да се правят заключения и прогнози. Моделите в медико-биологичните изследвания притежават генетични, физиологични и биохимични фактори и позволяват да се опишат основните закономерности на биологичните процеси. Детайлизацията и идентифицирането на тези модели от експериментални данни позволява описанието на реалните процеси в живи системи, проверка на техните механизми и прави тези модели евристични. Необходимостта от математическо и компютърно моделиране в медико-биологичните изследвания се обуславя от факта, че непосредственото изследване на много обекти е практически невъзможно или изисква много време, средства и специализирано техническо оборудване за експерименти.

Перспективите за развитие на математически модели в медико-биологичните изследвания почиват на използването на информационните технологии. Биоинформатика е прилагането на информационни и компютърни технологии за съхранение, организиране и анализиране на биологични данни.

Биологичната база данни е колекция от данни, която е организирана така, че съдържанието лесно може да бъде преглеждано, управлявано и актуализирано. Биологичните база данни са важен инструмент, подпомагащ учените да разберат множеството биологични явления. В дисертацията е използвана биологичната база данни Protein Data Bank Europe (PDBe) от където е взет теоретичен модел на делта опиоиден рецептор (DOR), използван при докинг процедурите в изследванията на лиганд-рецепторни взаимодействия.

Компютърното моделиране е всъщност математическо моделиране, при което се използват компютърни средства. Процеса на компютърно моделиране включва конструиране на модели, и тяхното

използване за решаване на задачи като анализ, изследване, оптимизация или синтез.

Съществен момент при математическия модел е степента на точност, респективно адекватността с реалния обект, която е в тясна зависимост с парадигмите в конкретната научна област (Mason R. L., R. F. Gunst, J. L. Hess, 2003). Проблемът за избор на „оптимален“ модел е един от най-основните проблеми в анализ на данни. Изследването за избор на математически модел е един актуален научен проблем, което е видно от големия брой световно извести учени, които работят по този проблем.

За намиране на индивидуалните „оптимални“ модели в класовете кандидат модели се използват различни методи за фитване, като метод на най-малките квадрати (LS), стабилна регресия (RR) и минимакс фитване.

На практика, изборът на модел и оценката са основни въпроси, като ключов аспект е изборът на „оптимален“ модел от набор от кандидат модели. Задачата за избора на модели които „най-добре“ описват данните се е превърнала в привидно повсеместна част от приложните изследвания. Ако приемем, че съществува един основен модел, който описва данните, критериите за моделен подбор се опитват да намерят този модел.

Често използвани в литературата, критерии за оценката на модели от различни класове са например информационен критерий на Акайке (AIC), информационен критерий на Бейс (BIC) и средно квадратично отклонение (RMSD). В информационно-теоретичния подход, препоръчан от (Akaike H., 1974; Kullback S., R. A. Leibler, 1951), информационното несъответствие се счита като основен критерий за оценка на добротата на модел като приближение до истинското разпределение, което генерира данните. AIC е получен като асимптотичната приблизителна оценка на Kullback-Leibler информационното несъответствие и предоставя полезен инструмент за оценка на модели по метода на максималната вероятност. На практика BIC, предложен от (Schwarz G., 1978) също е широко използван като критерий за избор на модел и може да се прилага за оценка на модели чрез метода на максималната вероятност. Въпреки това компромиса между добротата на фитване и простотата на модела при тези критерии не осигурява солидна основа за решаване на проблема с фитването на криви, тъй като тези критерии игнорират както статистическата адекватност, така и надеждността на изводите за проблема.

С разработването на гъвкави техники за моделиране, става необходимо разработването на критерии за избор на модели от различни класове, оценявани по методи, различни от метода на максималната вероятност. В дисертацията са предложени и минимизиране на максималния остатък (MMR) и критерий Хаусдорфово разстояние (HDC), като критерии за избор на модел и са сравнени с други известни в литературата критерии.

Настоящият дисертационен труд е разработен с помощта на специалисти от Центъра за функционални изследвания в спорта и кинезитерапията и Центъра за съвременни биоинформатични изследвания, към Югозападния Университет “Неофит Рилски” в Благоевград. Изследванията са проведени в Центъра за функционални изследвания в спорта и кинезитерапията към Факултета по обществено здраве и спорт на Югозападния Университет “Неофит Рилски” в Благоевград, където непрекъснато се провеждат подобен вид експерименти и се налага моделиране и оптимизация.

2. Цели и задачи на дисертацията

Цели на дисертационния труд:

Целите на настоящата дисертация са: (1) компютърно моделиране и избор на „оптимални“ модели с различни критерии за оптималност, при обекти свързани с медико-биологични изследвания и (2) разработване на подходящи софтуерни програми за рутинно приложение при тези изследвания.

Задачи:

1. Да се изследва и анализира съвременното състояние на моделирането и оптимизацията, както и тяхното приложение в областта на медико-биологични изследвания;

2. Да се проведат съответни медико-биологични изследвания с адекватен подход съобразно етиката на научните изследвания;

3. Да се регистрират динамометрични измервания с изокинетична динамометрична система при изометричен протокол при различни ъглови позиции и да се подберат подходящи методи и софтуер за статистическа обработка на получените експериментални данни;

4. Да се моделира торг-ъглова позиция зависимости на екстензорите на лактя;

5. Да се подберат подходящи математически методи и софтуер за оценка на параметрите на индивидуалните модели в различни класове модели при моделиране на торг-ъглова позиция зависимости;

6. Да се пробират подходящи известни, в литературата критерии за избор на „оптимален“ модел от различни класове модели, описващ торг-ъглова позиция зависимости;

7. Да се изследва възможността на Хаусдорфовото разстояние, като критерий за избор на „оптимален“ модел от различни класове модели при моделиране на биомеханични зависимости, като се сравни с други известни критерии;

8. Да се изследва възможността на минимизиране на максималния остатък, като критерий за избор на „оптимален“ модел от различни класове модели при моделиране на биомеханични зависимости и да се сравни с други критерии;

9. Да се разработи софтуерна програма за пресмятане на критериите за избор на „оптимален“ модел от различни класове модели;

10. Да се оценят параметрите в регресионното уравнение за изчисляване на максимална кислородна консумация (VO_{2max}) при нетренирани лица;

11. Да се подберат подходящи методи и софтуер за оценка на параметрите в регресионното уравнение за изчисляване на VO_{2max} за посочения контингент;

12. Да се разработи софтуерна програма за пресмятане на VO_{2max} с тест на Astrand-Ryhming, подходяща за рутинно прилагане;

13. Да се намери зависимост между резултати от докинг на селективни енкефалинови аналози и ДОР и *invitro* изследвания с корелационен анализ и да се оценят използваните софтуери за докинг;

14. Да се интерпретират получените резултати от проведените изследвания и да се формулират изводи.

3. Предмет на изследването

Предмет на изследването е избора и оценката на модели, използвани в медико-биологични изследвания като: зависимостта торг-ъглова позиция на екстензорите на лакътя; максимална кислородна консумация VO_{2max} ; и зависимост между резултати от докинг и биологични изследвания при лиганд-рецепторни взаимодействия.

4. Методи на изследване

За провеждане на динамометричните измервания и оценяването на максималната изометрична сила на екстензорите на лакътя, е използван изокинетичен динамометър Biodex.

За определяне на VO_{2max} с максимален тест, е проведен стъпаловидният максимален тест със спироергометрична система

QUARK CPET, Cosmed (USA, Italia) при натоварването с тредмил (h/P/COSMOS, Germany, модел Pulsar). За определяне на VO_{2max} със субмаксимален тест на Astrand-Ryhming, е проведен Astrand-Ryhming чрез натоварване на велоергометър при определена мощност, която е функция от възрастта, нивото на тренираност и телесната маса.

При докинг изследванията пептидите за които има данни от *in vitro* изследвания, са моделирани и оптимизирани с използването на софтуерния продукт Avogadro и подложени на докинг процедура с помощта на софтуерните продукти MVD и GOLD.

Обработката на експерименталните резултати посредством вариационен анализ (при изследванията на торг-ъглова позиция зависимости и VO_{2max}), корелационен анализ (при изследванията на лиганд-рецепторни взаимодействия) и непараметричен тест на Mann-Whitney (при изследванията на VO_{2max}) е направена с използването на софтуерния продукт GraphPad Prism.

При прилагането на регресионен анализ са използвани различни методи за фитване и софтуерни продукти. Използвани са метод на най-малките квадрати (LS), стабилна регресия (RR) и минимакс фитване. Използвани са софтуерните продукти GraphPad Prism, Matlab и SPSS Statistics.

За оценката и избора на “оптимален” модел от набор класове модели при моделирането на зависимости торг-ъглова позиция на екстензорите в лакътя са прилагани често използваните в литературата критерии за избор на модел: информационен критерий на Акайке (AIC), информационен критерий на Бейс (BIC) и Средно квадратично отклонение (RMSD), както и предложените в дисертацията критерии: минимизиране на максималния остатък (MMR) и критерий Хаусдорфово разстояние (HDC).

Възможностите на MMR и HDC като критерии за избор на „оптимален“ модел от различни класове модели са проверени, чрез сравняването им с критериите AIC, BIC и RMSD при биомеханични изследвания.

Програмните кодове на разработени в дисертацията програми и алгоритми са написани на езика C# с използването на средата за разработка Visual C # 2010 и библиотеката Framework 4.0.

5. Научна новост

Описана е последователна стратегия при анализа, математическото и компютърното моделиране, и избора на „оптимални“ модели описващи експериментални данни при медико-биологични изследвания, като са използвани различни техники за моделиране и критерии за оптимизация.

Намерен е „оптималния“ модел който може да помогне за обяснението на зависимостта торг-ъглова позиция на екстензори на лакътя при жени и може да се използва успешно при прогнозирането на тези отношения, като спестява време и ресурси.

Намерени са „оптималните“ стойности на параметрите в регресионното уравнение за изчисляване на VO_{2max} с тест на Astrand-Ryhming при нетренирани лица и е разработена софтуерна програма за изчисляването на VO_{2max} с този субмаксимален тест.

Изследвана е възможността за използване на MMR и на Хаусдорфовото разстояние (HD) като критерий HDC за избор на „оптимален“ модел на биомеханични данни за зависимости торг-ъглова позиция чрез сравняването им с често използваните в литературата критерии AIC, BIC и RMSD и е установено, че MMR и HDC може успешно да се прилагат като критерий за избор на „оптимален“ модел в този вид биомеханични изследвания. Разработени са също и софтуерни програми за изчисляване на критериите за избор на „оптимален“ модел.

6. Апробация

Резултатите от дисертационния труд са обсъждани и докладвани на следните научни прояви:

- International Scientific Conference – FMNS2011, 2011, Blagoevgrad, Bulgaria;
- 14th National Biostatistics Congress with International Participation, 2012, Kadir Has Center, Kayseri, Turkey;
- International Scientific Conference „Functional studies in sport and kinesitherapy“, 2014, Blagoevgrad, Bulgaria;
- International Scientific Conference „Enhancement of the competitiveness and restructuring of the food subsectors through benchmarking“, 2014, Blagoevgrad, Bulgaria.

7. Публикации

Получените резултати са публикувани в 3 статии в списания (от които 1 самостоятелна в списание Доклади на БАН), и 4 в доклади (от които 1 самостоятелна) на международни конференции. 1 статия е подготвена за публикуване (In Preparation). 7 публикации са на английски език и 1 е на български език.

СЪДЪРЖАНИЕ НА ДИСЕРТАЦИОННИЯ ТРУД

Настоящата дисертация съдържа 140 страници и се състои от използвани съкращения, увод, три глави, изводи, списък на цитираната литература и седем приложения. Включени са 30 фигури и 14 таблици. Цитираната литература обхваща 192 заглавия, от които 14 на кирилица и 178 на латиница.

ПЪРВА ГЛАВА: ПРОБЛЕМАТИКА И ОБОСНОВКА НА ИЗСЛЕДВАНИЯТА

Първа глава съдържа основните аспекти на обектите моделирани в изследванията и мотивировката на изследванията, приложимостта и ролята на математическото и компютърното моделиране при този вид изследвания, основните методи за моделиране на зависимости и критерии за оценка на модели, както и подбора на съответния софтуер.

1.1. Биологични основи на обектите моделирани в изследванията

В тази част на първа глава са дадени основните аспекти на моделираните в изследванията обектите: зависимости торг-ъглова позиция, максимална кислородна консумация VO_{2max} и лиганд-рецепторни взаимодействия.

1.1.1. Биомеханични аспекти на скелетно-мускулна активност. Зависимост торг-ъглова позиция

Биомеханиката е интердисциплинарна наука, която изследва активните движения на живите организми, като от особен интерес са движенията на човека и локомоцията, т.е. придвижването в пространството. В локомоцията участват предимно опорно-двигателната система, в изграждането на която влизат костите, свързани помежду си чрез стави в скелет и мускулите.

Костите и ставите са пасивни елементи на опорно-двигателния апарат на човека, а скелетните мускули - активната. Костно-ставните сегменти играят роля на анатомични лостове, чрез които се мултиплицира действието на мускулите и се печели сила. Понеже движението в ставите е въртливо, то силата на мускулите се измерва като въртящ момент или торг в mN , при това на група мускули.

Точното определяне на точката на приложение на мускулната сила върху анатомичните лостове определя и свойствата им да генерират сила. Ъгълът на приложение на мускулната сила и приложната ѝ точка са определящи за движението, което ще се извърши в дадена става (Флойд Р. Т., 2008). Промяната на ъгъла в ставата води до промяна в изходната дължина на мускулите, които участват в движението на дадена става.

Зависимост торг-ъглова позиция. Съотношението между вектора на ротаторната и стабилизиращата компонента се променя в зависимост от изходната ъглова позиция в ставата. Следователно силата, която един мускул може да генерира при своята контракция не е постоянна величина, а зависи от ставния ъгъл.

Зависимостта торг-ъглова позиция е обект на изследователски интерес и математическо моделиране при различни по характер изследвания като: неврофизиологични и биомеханични (Enoka R.M., 2002; Pencheva N., M. Kokova, S. Dencheva, 2013; Pinter I., B. Maarten, S. Jeroen, 2010), клинични, приложни изследвания в спорта, ергономията, кинезиологията (Zatsiorsky V.M., Prilutsky B.I., 2012).

Кинезиология на мускулното действие в лакътната става. Мускулното действие в лакътният комплекс, дава възможност, както за движение, така и за динамична стабилизация при задържане на ставата в дадена позиция (Попов Н., 2003). Мускулите в лакътя осъществяват движения само в сагитална равнина, които са: флексия или сгъване и екстензия или разгъване. При флексията ъгълът между костите, формиращи ставата намалява, а при екстензията се увеличава.

Изокинетичната динамометрия. Изокинетичната динамометрия измерва въртящия момент (торг) и по този начин сумарната сила на група мускули, при различен режим на съкращение, а чрез подходящи допълнителни параметри оценява и двигателните качества, бързина и издръжливост (Пенчева Н., 2010а). Изокинетичната динамометрия се използва като референтен метод за оценка на сумарен торг.

Изследването на (Pinter I., B. Maarten, S. Jeroen, 2010) е насочено към връзката между изометричната сила на флексорите и екстензорите, при това в комбинация с ЕМГ активност на biceps и triceps brachii, но измерванията са проведени не в сагиталната, а в трансверзалната равнина с оглед да се елиминира ефектът на гравитационната сила и да се потърси връзка с движението в раменната става, при различни нива на повдигане на мишницата. Макар и при различен дизайн, резултатите от това изследване до голяма степен провокираха настоящото.

1.1.2. Тестове за измерване и изчисляване на максимална кислородна консумация

Максималната кислородна консумация (VO_{2max}) е важен показател за аеробната работоспособност и издръжливостта на човека при физическо натоварване (Astrand P-O, 2002; Пенчева Н., 2010б). Тя представлява максималното количество кислород, което организма може да усвои за една минута, като вентилира въздух и затова се оценява чрез тестове за физическо натоварване, които могат да бъдат

различни, в зависимост от изследваните лица, апаратурни възможности, както и методи за моделиране на експерименталните резултати (Душков В., Стефанова Д., Джарова Т., 1986; Пенчева Н., Коцев Ч., Манчев А., 2010; Манчев А., 2014).

Показателят VO_{2max} е обект на изследване в настоящата разработка, защото рутинните методи за неговото определяне в широката практика се провеждат при субмаксимално натоварване, което е безопасно и достъпно, посредством математическо моделиране на експерименталните данни, на базата на известни в литературата регресионни уравнения. Така се избягват скъпо струващи методи, които изискват максимално натоварване на лицата до отказ, сложни системи за газов анализ и други.

1.1.3. Лиганд-рецепторни взаимодействия

Традиционно, лекарствени средства се откриват от синтезирани съединения в резултат от многоетапни процеси. Тези съединения след това се подлагат на серия от *in vivo* или *in vitro* тестове на експериментални животни и хора-доброволци. Много биологични процеси, важни за функционирането и преживяването на клетките (и организмите) се регулират от биологично активни вещества като хормони, невротрансмитери и други сигнални молекули, които се наричат с общото название лиганди. Такава регулация се осъществява чрез взаимодействие на тези природни молекули с рецептори.

Опиоидните пептиди (късоверижни пептиди) и техните рецептори са част от неврохормоналната система. Ендогенните опиоидни пептиди са малки молекули, които се произвеждат в централната нервна система (ЦНС) и различни жлези в тялото.

В настоящото изследване е направен опит да се намери корелация между данните получени от *in vitro* тестовете и докинг изследвания.

1.2. Математическо и компютърно моделиране в медико-биологичните изследвания

В тази част на първа глава са дадени понятието за математически и компютърен модел, ролята на моделирането, съставянето на модели и формулирането на набор от кандидат модели.

1.2.1. Понятие за математически и компютърни модели

Експериментални данни, математически и компютърни модели на данни се използват в емпиричните науки, за да се направят изводи относно процеси и параметри представляващи интерес. (Box G.E.P., W.G. Hunter, J.S. Hunter, 1978; Lunneborg C.E., 1994; Shenk T.M., A.B. Franklin, 2001). Математическото и компютърното моделиране са

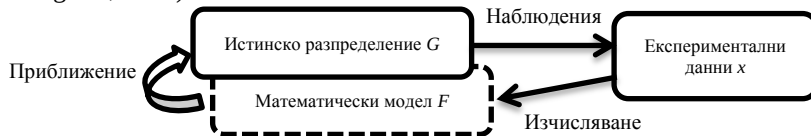
ключови въпроси при анализа на данни в медико-биологичните изследвания.

Математическите модели се развиват от уравнения, които определят как една система се променя от едно състояние в следващите (диференциални уравнения) и/или как една променлива зависи от стойността, или състоянието на други променливи (уравнения).

Компютърното моделиране като един съвременен метод за научни изследвания се основава на построяването на математически модели за описание на изучаваните процеси чрез използване на компютри.

1.2.2. Роля на математическите и компютърните модели

Математическите и компютърните модели играят важна роля при анализа на експериментални данни. След като идентифицираме даден модел, може да направим различни форми на заключения като прогноза, контрол, извличане на информация, откриване на знания, валидиране, оценка на риска или вземане на решения. Целта на математическото и компютърното моделиране е да изгради модел, който възможно най-точно се доближава до истинската структура, чрез използване на наличните данни (Фигура 1.10.) (Sadanori K., G. Kitagawa, 2008).



Фигура 1.10. Оценка на правилната структура на основата на математически модели. *Адаптация по* (Sadanori K., G. Kitagawa, 2008).

1.2.3. Съставяне на математически и компютърни модели.

Формулиране на множества от кандидат модели

Когато съставяме математически модел, обикновено се прави опит да се определи дали експериментални данни са в съответствие с конкретна теоретична връзка и да се намери модел в даден клас модел M описващ тази връзка. Предполага се, че експерименталните данни се състоят от набор от n стойности на някоя измерена променлива y : $y_1, \dots, y_n$, съответстваща на n стойности на някоя независима променлива x : $x_1, \dots, x_n$. Така се получават точки (x_i, y_i) , за $i = 1, \dots, n$, образувачи множество от точки с експерименталните данни $A = \{(x_1, y_1), \dots, (x_n, y_n)\}$. Множеството от точки A може да бъде моделирано с математическа функция $y = f(x; p_1, \dots, p_m)$, с променлива x и

параметри $p_1, \dots, p_m$, чрез използване на подхода за фитване на криви и компютърна програма за фитване и да получим компютърен модел (Sadanori K., G. Kitagawa, 2008).

Формулиране на множество от кандидат модели

Докато използваме емпирични данни за анализ, е необичайно моделът да бъде известен. Трябва да се разгледат редица алтернативни форми на модел, както и специфични обяснителни променливи, които се използват при моделирането на зависимата променлива.

Формулирането на множество от кандидат модели е частично субективно изкуство. За изготвянето на множество от априорни кандидат модели можем да се използва публикуваната литература и опита в съответните науки.

Използвани математически и компютърни модели в медико-биологичните науки

В медико-биологичните науки математическите и компютърните моделите са полезни за разбиране на структурата на системи, изчисляване на параметри представляващи интерес и свързаните с тях вариация – ковариационна матрица, предсказване на резултати и отговори, и тестване на научни хипотези.

За изготвянето на множество от априорни кандидат модели в настоящата дисертация се използва публикуваната литература и опита при изследвания на моделираните обекти в дисертацията.

Зависимости торг-тъглова позиция. Днес, изследването на връзката между максимален изометричен торг (въртящ момент) и ъглова позиция е обект на засилен изследователски интерес (Pinter I., B. Maarten, S. Jeroen, 2010; Bober T., Kulig K., Burnfield J., Pietraszewsk B., 2002; Wilson E., E. Rustighi, P. Newland, B. Mace, 2012), защото тя повишава знанията и способността да се оцени производство на максимален торг през цялата гама от движения. Традиционно, моделирането на въртящ момент-ъгъл зависимости се прави с класове от втора, трета и четвърта степен полиноми (Pinter I., B. Maarten, S. Jeroen, 2010; Brughelli M., J. Cronin, K. Nosaka, 2010; Pencheva N., M. Kokova, S. Dencheva, 2013). В техните изследвания, при оценката на класовете модели са използвани критерий R^2 и AIC.

При изследванията в дисертацията фитването на експерименталните данни за зависимост торг-тъглова позиция, е направено с полиноми от втора, трета и четвърта степен и е избран „оптимален“ модел с помощта на различни методи и критерии за избор на „оптимален“ модел.

Максимална кислород консумация (VO_{2max}). Максимална кислород консумация (VO_{2max}), известна също като функционален

аеробен капацитет, представлява максималната степен на използване на кислород чрез упражняване на мускулите и се счита за „златен стандарт“ мярка за функционалната граница на сърдечната и система (Saltin B, S.O. Strange, 1991; Siconolfi S.F, C.E. Garber, T.M. Lasater, R.A. Carleton, 1985). Нейното оценяване става чрез максимални или субмаксимални тестове. Поради трудното прилагане на максималните тестове за масови изследвания, в практиката са прилагани много често субмаксимални тестове, които са достатъчно надеждни и достъпни за практиката на спортния педагог. Такъв например е теста на Astrand-Ryhming за индиректно определяне на VO_{2max} (Adams G., 1998).

В дисертацията е направено моделиране с регресионно уравнение и оценка параметрите на известния модел на Adams за определяне на VO_{2max} с тест на Astrand-Ryhming, при нетренирани лица.

Лиганд-рецепторни взаимодействия. В резултат на биологичните изследвания се получават данни за различни параметри: IC_{50} – потентност, K_A – афинитет и e_{rel} – ефикасност. Тези параметри са указание за биологичното действие на съответния лиганд по отношение на дадения рецептор. IC_{50} е концентрацията, необходима за 50 % инхибиране на ефекта на лиганда. K_A и e_{rel} стойностите се изчисляват с експлицитни формули (Pencheva N., M. Kokova, S. Dencheva, 2013).

Емпиричните данни за връзката между ефекта (отговора), предизвикан от лекарство и неговата концентрация, изразена логаритмично, могат да се моделират с уравнението:

$$E = \frac{M[A]^P}{[A]^P + [EC_{50}]^P}, \quad (1.3)$$

където E е отговор, M е максимален отговор, $[EC_{50}]$ – концентрация, предизвикваща половината от максималния отговор, а P – коефициент на Hill за кривата концентрация-отговор (Barlow R. B., S. M. Bond, C. Grant, D. S. McQueen, Z. Yaqoob, 2001).

(Barlow R.B., 1975; Barlow R.B., S.M. Bond, E. Bream, L. Macfarlane, D.S. McQueen, 1997; Barlow R. B., S. M. Bond, C. Grant, D. S. McQueen, Z. Yaqoob, 2001) изследват криви за физиологични антагонисти при изотонични и изометрични измервания. Взаимодействието на двата вида на лекарството е изследвано чрез получаване на концентрация-ефект криви за агонисти в присъствието на нарастващи концентрации на антагонисти. Експерименталните резултати доказват, че голям процент от съответстващи си криви в разгледаните случаи са хиперболични.

1.3. Методи и критерии за оптималност при моделиране на зависимости

В тази част на първа глава е представена последователността при моделирането на зависимости, представени са основни методи за фитване и известни критерии за избор на „оптимален“ модел от различни класове модели.

1.3.1. Методи и критерии за намиране на оптимален модел в даден клас

За намиране на индивидуален „оптимален“ модел в даден клас модели чрез оценка на параметрите на този модел съгласно даден критерий, се използват различни фитващите методи.

Най-малките квадрати (LS) за линейни модели са сравнително лесни за изчисляване и затова те са познати с ранната история на тяхното използване (Weisberg S., 1985). За разлика от това вероятностните методи на Fisher, често изискват итеративен числени методи и по този начин не са популярни преди широко разпространената наличност на персонални компютри и развитието на лесен за използване софтуер. Например проста линейна регресия, където променлива отговор (y) се моделира като линейна функция на обяснителна променлива (x), като $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$. Тук ε_i са грешки (остатъци), които често са моделирани като независими нормални случайни величини със средна стойност 0 и постоянна дисперсия σ^2 . LS оценки на β_0 и β_1 са тези, които намаляват до минимум $\sum (\varepsilon_i)^2$ - оттам идва и името най-малките квадрати.

Вероятностните ML методи са много по-общи, и малко по-трудно за разбиране. Нека например, при многократно повтаряне на експеримент са получени стойности: $x_1, \dots, x_n$ на случайната променлива X . Приемаме, че X е разпределена с плътност $f(X, a_1, \dots, a_k)$, където $a_1, \dots, a_k$ са неизвестни параметри. Методът на максималната вероятност се състои в избор на такива стойности на $a_1, \dots, a_k$, при които редицата $x_1, \dots, x_n$ се реализира с максимална вероятност.

Известен е също и минимакс алгоритъм за апроксимация или L^∞ апроксимация (Powell M. J. D., 1981; Van den Bos A., 1988), който представлява метод за намиране на приближение, което намалява максималната грешка. Например, за да се сближи функция $e^{f(x)}$ с функция $p(x)$ в интервала $[a, b]$ алгоритмът минимакс ще намери функция $p(x)$, която ще се сведе до минимум:

$$\max_{a \leq x \leq b} [f(x) - p(x)] \quad (1.6)$$

1.3.2. Критерии за оценка и сравнение на оптимални модели от различни класове

Критериите за избор на модел от различни класове модели може да се използват да се оценят и сравнят всеки един от индивидуалните „оптимални“ модели в дадени класове.

В теория на информация известното Kullback-Leibler „разстояние“ между два модела (всъщност, вероятностни разпределения), представлява основна величина в областта на науката (Burnham K. P., D. R. Anderson, 2002). През 1973 г. Akaike получил оценител на очакваното Kullback-Leibler разстояние въз основа на максимизиране на логаритмичната вероятност на Fisher (Akaike H., 1973). Неговата мярка, сега се нарича информационен критерий на Акайке (AIC), и е нова парадигма за избор модел при анализа на емпирични данни. В общия случай, $AIC = -2 * \log(L) + 2k$, където k е броят на параметрите в модела, а L е максималната стойност на вероятностната функция за разглеждания модел.

Средно квадратично отклонение (RMSD) или средна квадратична грешка (RMSE) е често използвана мярка за разликите между стойностите предсказани от един модел и действително наблюдаваните стойности. RMSD е добра мярка за точност, при сравнение на различни модели (Hyndman R. J., A. B. Koehler, 2006; Armstrong J. S., F. Collopy, 1992; Myung I. J., 2000). RMSD на модела по отношение на параметрите се определя като корен квадратен от средна квадратична грешка.

Bayesian изследователи са взели малко по-различни подходи и предположения, и са предложили няколко алтернативни метода за избор на модел, като например и информационен критерий на Бейс (BIC): $BIC = -2 * \log(L) + k * \ln(n)$, където k е броят на параметрите в модела, а L е максималната стойност на вероятностната функция за разглеждания модел а n е размера на извадката (Schwarz G., 1978).

Процедурите от типа на AIC не предоставят задоволително решение на проблема за избора на фитващи криви, защото те не се отнасят изключително важно към адекватност на въпроса кога фитващата крива улавя закономерности в данните.

1.3.3. Хаусдорфовото разстояние като критерии за оптималност

Хаусдорфовото разстояние (HD) е математическа конструкция за измерване на близостта на две множества от точки, които са подгрупи на метричното пространство. Известно е използването на Хаусдорфовото разстояние в много различни области: разпознаване на образи (Daniel P., A. Gregory, J. William, 1993), съвпадение на бинарни

изображения (Shapiro M., M. Blaschko, 2004), повърхностно съвпадение (Manocha D., C. Erikson, 1998), сходство на две траектории (Nutanong S., E. H. Jacox, H. Samet, 2011).

Всичко това провокира голям интерес да се приложи Хаусдорфовото разстояние като критерии за избор на „оптимален“ модел от различни класове модели и да се оцени неговата ефективност, като се сравни с други често използвани критерий за избор на „оптимален“ модел намерени в литературата.

1.3.4. Минимизиране на максималния остатък като критерии за оптималност

Идеята за прилагането на минимизиране на максималния остатък (MMR), като критерий за избор на „оптимален“ модел идва от минимакс метода за фитване. При този метод за фитване, намираме индивидуален „оптимален“ модел в избран клас модели чрез минимизиране на максималната грешка. В дисертацията е приложен този метод и за сравнение на избраните индивидуални „оптимални“ модели в отделните класове модели, намерени с различни методи за фитване.

1.3.5. Подбор на софтуер

Като инструмент за изследване, симулиране, визуализация на резултатите от проведените изследвания в дисертацията, както и за разработването на софтуерни програми е използван подходящ лицензиран софтуер в университета и пробни (trail) версии.

GraphPad Prism 6.0. За обработка на резултатите от изследванията: вариационен анализ, корелационен анализ, непараметрични тестове и еднофакторен регресионен анализ е използван софтуерен продукт GraphPad Prism.

Matlab 2009. Тъй като в GraphPad Prism няма възможност за нелинейна регресия използваща фитване по метода минимакс, в този случай е използван софтуерния продукт Matlab.

SPSS Statistics 22.0 (Free trial). При множествен (многофакторен) регресионен анализ когато имаме две или повече независими променливи е използван софтуерен продукт SPSS Statistics, тъй като GraphPad Prism няма възможност за изпълняване на този вид регресионен анализ.

Visual C # 2010 Express. За разработването на софтуерните програми използвани в дисертацията е използвана средата за разработка Visual C # 2010 и библиотеката Framework 4.0.

1.4. Моделиране на лиганд-рецепторни взаимодействия

През последните десетилетия все повече се развиват, така наречените *in silico* методи, които използват изчислителната среда

като експериментална лаборатория. Терминът *in silico* се отнася до използване на компютри за имитиране на биологични обекти и процеси. По-голяма част от *in silico* методите се използват паралелно с *in vivo* и *in vitro* изследванията с цел прецизно моделиране и валидиране на широк спектър от приложения при дизайна на лекарства и при оптимизиране на характеризирането на фундаментални фармакологични свойства на молекулите на лекарствата (Ekins S., J. Mestres, B. Testa, 2007).

Съществуват различни методи при *in silico* лекарствения дизайн (Wadood A., N. Ahmed, L. Shah, A. Ahmad, H. Hassan, S. Shams, 2013), като хомоложно моделиране, виртуален високо-ефективен скрининг, количествена връзка структура – активност (QSAR), конформационен анализ, молекулно динамични симулации, Молекулен докинг и други.

Молекулен докинг използван в нашите изследвания е техника, която предвижда благоприятната ориентация на една молекула към друга, когато се свързват и образуват стабилен комплекс.

Софтуер за моделиране на лиганд-рецепторни взаимодействия

GOLD 5.2. Той се предлага като част от GOLD Suite, които включва два допълнителни компонентни софтуера, Hermes и GoldMine. В направените изследвания в дисертацията е използвана фитнес функцията *GoldScore*, представена:

$$GOLDScore = S_{hb_ext} + S_{vdw_ext} + S_{hb_int} + S_{vdw_int}, \quad (1.8)$$

където S_{hb_ext} е енергията на външните водородни връзки между протеина и лиганда, S_{vdw_ext} е енергията на външните ван дер Ваалсови взаимодействия между лиганда и протеина, S_{hb_int} е енергия на вътрешномолекулните водородни връзки в лиганда и S_{vdw_int} е енергия на вътрешномолекулните деформации в лиганда.

Molegro Virtual Docker 6.0 (MVD) (Thomsen R., M. H. Christensen, 2006). MVD е интегрирана платформа за прогнозиране на протеин - лигандни взаимодействия. Основната оценъчна функция за ефикасността на докинга е E_{score} функцията, която се представя със следния израз:

$$E_{score} = E_{inter} + E_{intra}, \quad (1.9)$$

където E_{score} е оценъчната функция от докинг, E_{inter} – енергия на лиганд-протеиновото взаимодействие и E_{intra} – вътрешна енергия на лиганда.

Avogadro 1.1.0 (<http://avogadro.openmolecules.net/>). Avogadro е усъвършенстван молекулен редактор и визуализатор. Програмата е използвана за моделиране и оптимизиране на структурите на лигандите.

ВТОРА ГЛАВА: МЕТОДИЧНИ ПОДХОДИ

Втора глава съдържа описание на използваните методични подходи при медико-биологични изследвания и моделирането на зависимости. Представена е методологията на изследванията, статистическата обработка на експерименталните резултати, математическото и компютърно моделиране на зависимостите и използваните критерии за избор на „оптимален“ модел.

2.1. Динамометрични измервания на екстензорите на лакътя с изокинетичен динамометър

За получаване на експериментални данни необходим за моделиране на торг-ъглова позиция зависимости на екстензорите на лакътя при жени, са проведени динамометрични изследвания. Изследванията са проведени в Центъра за функционални изследвания в спорта и кинезитерапията към Факултета по обществено здраве и спорт на Югозападния Университет “Неофит Рилски” в Благоевград. Динамометричните параметри се регистрираха върху доминантната ръка. Изследваните лица бяха жени, студентки от Югозападен Университет “Неофит Рилски”, Благоевград, от специалностите “Кинезитерапия” и Кинезиология”. Сумарният въртящ момент или торгът на екстензорите на лакътя на дясната ръка беше измерен в Nm по време на максимална, волева, изометрична контракции (две повторения с продължителност от 3 s, *като се отчита по-високата стойност*, разделени от 10 s релаксация) за всяка една от следните 10 ъглови позиции в ставата: 15°, 30°, 45°, 60°, 70°, 90°, 105°, 120°, 135° и 150°.

2.2. Максимален тест и тест на Astrand-Ryhming за определяне на максималната кислородна консумация (VO_{2max})

Експерименталните данни необходим за оптимизиране стойностите на параметрите в регресионното уравнение (Adams G., 1998) за изчисляване на VO_{2max} за конкретен контингент от студенти са получени в резултат от провеждането на максимален спироергометричен тест за директно измерване на VO_{2max} и тест на Astrand-Ryhming за индиректно определяне на VO_{2max} . При оптимизация на параметрите в регресионното уравнение зависимата променлива VO_{2max} е получена от експериментите с максимален тест, а стойностите на независимите променливи в уравнението (мощността на натоварването, възраст и сърдечната честота в състояние на steady state) са получени от експериментите с тест на Astrand-Ryhming.

Максимален, спироергометричен, стъпаловиден тест за определяне на VO_{2max} . При настоящото изследване стъпаловидният максимален тест се проведе съвместно с екипа на Центъра за

функционални изследвания в спорта и Кинезитерапията при ЮЗУ „Неофит Рилски” със спироергометрична система QUARK CPET, Cosmed (USA, Italia) при натоварването с тредмил (h/P/COSMOS, Germany, модел Pulsar).

Изследвани са 19 мъже, неактивно занимаващи се със спорт студенти на възраст 19-25 години (21.0 ± 2.2 години) от различни специалности на Университета, по протокол на Илиев (1980). Относно натоварването, то се изразяваше в следното: - в първите 30 секунди от началото на теста изследваното лице изчакваше, за да бъдат регистрирани изследваните параметри; - на 30-та секунда започваше да се движи със скорост 5 km/h (ходене) и постепенно (стъпаловидно), през равни интервали от време (на 1.5 min) скоростта се покачваше с 1.2 km/h.

Субмаксимален тест на Astrand-Ryhmig за определяне на VO_{2max} . Astrand-ryhmig тест се изпълнява чрез натоварване на велоергометър при определена мощност, която е функция от възрастта, нивото на тренираност и телесната маса. До 5-та, 6-та мин от натоварването се получава т. нар. ергостаза, поради което по отчетената пулсова честота в края на 6-та минута от натоварването и по стойността на използваната мощ (W_{max}) се изчислява VO_{2max} . Така, индиректно се определя VO_{2max} . За достигане на относително устойчиво състояние на пулсовата честота (steady state) при 4-5-та минута от началото на натоварването бяха предварително проведени тестове за уточняване на коефициента, по който се пресмята натоварването. Стойностите на VO_{2max} се изчисляват със следното регресионно уравнение за мъже (Adams G., 1998):

$$VO_{2max} = ((P * 0.012) + 0.3) * \frac{(220 - age) - 61}{HR_{sm} - 61} \quad (2.1)$$

където:

- P е мощността на натоварването измерена във ватове (W); - age е възрастта на изследваните лица (години); - HR_{ss} е сърдечната честота (heart rate) в състояние на steady state ($y\delta/min$); - 0.012 и 0.3 са параметри в уравнението по литературни данни.

2.3. Докинг изследвания на лиганд-рецепторни взаимодействия

Структурите на 11 пептида: DPDPE, енкефалини и техни аналози за които има данни от *in vitro* изследвания (Pencheva N., P. Milanov, L. Vezenkov, T. Rajpanova, E. Naydenova., 2004), бяха моделирани и оптимизирани с използването на Avogadro и подложени на докинг процедура с помощта на MVD и GOLD 5.2. За целите на докинга бе

използван теоретичния модел на δ -опиоидния рецептор, публикуван в биологичната база данни PDBe id: 1o2c.

2.4. Вариационен анализ

При изследванията направени в дисертацията е използван вариационен анализ със софтуерния продукт GraphPad Prism за изчисляване на: - средна аритметична величина $\bar{X}$; - стандартното отклонение (SD); - стандартна грешка на средната (SEM); - коефициентът на вариация (V).

Вариационен анализ е приложен при статистически обработки на експериментални данни от измерване на зависимост торг-ъглова позиция в лакътна става и статистически обработки на експериментални данни от тестове за определяне на VO_{2max} .

2.5. Корелационен анализ

За намиране на корелация, между данните получени от докинг изследванията и биологичните изследвания е използван GraphPad Prism. Изчислени са коефициентът на обикновена линейна корелация на Пирсън (r) и коефициентът на рангова корелация на Спирмън (r_s).

2.6. Непараметричен тест на Mann-Whitney

При изследванията за VO_{2max} , за сравняване на средни при 2 групи е използван непараметричен тест на Mann-Whitney (за независими извадки) със софтуерния продукт GraphPad Prism. Теста на Mann-Whitney наричан още тест на ранговите суми е непараметричен тест, който сравнява средните рангове на две независими извадки.

2.7. Регресионен анализ и методи за намиране на оптимален модел в даден клас

При математическото и компютърното моделиране на торг-ъглова позиция зависимости с различни класове модели, както и при оценка на параметрите в регресионното уравнение за изчисляване на VO_{2max} , във връзка с направените изследвания в дисертацията, е използван регресионен анализ. Регресионен анализ е метод за изследване на зависимостите между зависима и множество независими променливи. Използвани са изброените по-долу методи за фитване.

Метод на най-малките квадрати. При нелинейната регресия, като и при линейната регресия, се приема, че разсейването на данните около „оптималната“ крива следва нормално разпределение. Това предположение води до целта на метода на най-малките квадрати: свеждане до минимум на сумата от квадратите на вертикалните отклонения или (разстоянията между Y -стойностите) между точките и кривата (Konishi S., G. Kitagawa, 2008). В случай на полиномни модели каквито са използвани в дисертацията при моделиране на торг-ъглова

позиция зависимости, критерия за оптималност по отношение на параметрите на този метод е минимизиране на сумата:

$$\sum_{\alpha=1}^n (y_{\alpha} - \hat{y}_{\alpha})^2 = \sum_{\alpha=1}^n \{y_{\alpha} - (\beta_0 + \beta_1 x_{\alpha} + \dots + \beta_m x_{\alpha}^m)\}^2, \quad (2.6)$$

където y_{α} е измерената стойност на зависимата променлива, а $\hat{y}_{\alpha}$ е изчислената стойност с полином от степен m .

Стабилна регресия. Експериментални грешки може да доведат до погрешни данни, чиито стойности са твърде високи или твърде ниски. Дори една единствена грешна стойност може да окаже влияние върху изчисляването на сумата на най-малките квадрати, и да доведе до подвеждащи резултати. Един от начините за справяне с този проблем е да се направи стабилна регресия, като се използва метод, който не е много чувствителен към нарушения на допускането за нормално разпределение на данните. При стабилна регресия извънредните стойности имат малко влияние (Motulsky H., R. Brown, 2006). Стабилна регресия е използвана при намиране на индивидуалните „оптимални“ модели в класовете модели, при моделиране на торг-ъглова позиция зависимости при изследванията в дисертацията.

Минимакс фитване. Проблемът минимакс също е често използван при фитване на криви (Van den Bos A., 1988). Тази процедура намалява максималният остатък (виж формула 1.6) (Rabinowitz P., 1968). Този метод е използван при моделиране на торг-ъглова позиция зависимости.

2.7.1. Регресионен анализ при моделиране на биомеханични зависимости

В случая на моделиране на биомеханични зависимости от измерването на торг-ъглова позиция в лакътната става, е направен еднофакторен регресионен анализ със софтуерните продукти GraphPad Prism и Matlab за определяне на зависимостта между зависимата променлива (торг) и независимата променлива (ъглова позиция) с различни фитващи методи и различни класове фитващи модели.

2.7.2. Регресионен анализ при моделиране на уравнението за VO_{2max}

При на моделиране с регресионното уравнение за определяне на VO_{2max} е направен многофакторен регресионен анализ по метод на най-малките квадрати със софтуерния продукт SPSS Statistics за определяна на параметрите в уравнението.

Оптимизационна процедура за тестване на локален или глобален екстремум

В използваните в дисертацията софтуерни пакети (GraphPad Prism, Matlab и SPSS) за нелинейна регресия няма пълна гаранция за това, че намереното решение е глобален екстремум. В софтуерните пакети, за да има висока степен на вероятност решението да е глобалния минимум, се препоръчва да се стартира с голям брой различни начални стойности на параметрите $a_i, i = 1, \dots, k$. Следвайки тези указания, при намирането на решенията, указани в дисертацията, са следвани тези препоръки. По-точно: ако намереното решение е $(a_1^{\text{opt}}, \dots, a_k^{\text{opt}})$, то отново е стартирана процедурата от точки $(a_1^{\text{opt}} \pm \varepsilon_1, \dots, a_k^{\text{opt}} \pm \varepsilon_k)$, където $\varepsilon_i \in [0, 200], i = 1, \dots, k$ и ε_i са достатъчно големи, и е получавано отново $(a_1^{\text{opt}}, \dots, a_k^{\text{opt}})$ за решение.

2.8. Критерии за избор на оптимален модел от различни класове

При избора на „оптимален“ модел от различни класове модели за определяне на зависимостта торг-ъглова позиция на екстензорите, във връзка с направените биомеханични изследвания в дисертацията са използвани AIC, BIC, RMSD, MMR и HDC, като критерии за оценка на моделите.

Информационен критерий на Акайке (AIC). Идеята на AIC е да избере модела, който минимизира вероятността за допускане на грешка, санкционирайки увеличаването на броя параметрите в модела (Acquah de-Graft H., 2010; Joseph B., L. Nicole, 2004; Konishi S., G. Kitagawa, 2008; Zucchini W., 2000; Akaike H., 1974).

Формулата използвана в дисертацията за изчисляване на този критерий е:

$$AIC = \begin{cases} n * \ln\left(\frac{RSS}{n}\right) + 2 * k, & \frac{n}{k} \geq 40 \\ n * \ln\left(\frac{RSS}{n}\right) + 2 * k + \frac{2 * k * (k + 1)}{n - k - 1}, & \frac{n}{k} < 40, \end{cases} \quad (2.9)$$

където:

n е броят на точките с експериментални данни; k е броят на параметрите в модела плюс един (тъй като регресия е "оценка" на сумата от квадратите, както и на стойностите на параметрите); RSS или остатъчна сума от квадрати е сумата от квадратите на вертикални отклонения от всяка точка на данните до „оптималната“ фитващата крива.

Информационен критерий на Бейс (BIC). BIC има най-голям постериорна вероятността (Acquah de-Graft H., 2010; Joseph B., L.

Nicole, 2004; Konishi S., G. Kitagawa, 2008). Модела който минимизира този критерий се избира за „оптимален“.

Формулата използвана в дисертацията за изчисляване на този критерий е:

$$\text{BIC} = n * \ln\left(\frac{\text{RSS}}{n}\right) + k * \ln(n) \quad (2.12)$$

Средно квадратично отклонение (RMSD). Формулата за изчисляване на RMSD заложена в софтуерния продукт GraphPad Prism и използвана в дисертацията е:

$$\text{RMSD} = \sqrt{\frac{\text{RSS}}{n - k}}, \quad (2.13)$$

където n е броят на точките с експериментални данни; k е броят на параметрите в модела; RSS е сборът от квадратите на вертикални отклонения от всяка точка на данните до „оптималната“ фитващата крива.

Минимизиране на максималния остатък (MMR). MMR е въведен като критерий за избор на „оптимален“ модел в направените биомеханични изследвания за определяне на зависимостта торг-ъглова позиция на екстензорите.

За намерените от съответните фитващи методи, индивидуални „оптимални“ модели $P^*(M_j)$ в избраните класове M_j , $j = 1, \dots, \kappa$, за всеки от фитващите методи е минимизиран максималния остатък между измерените и изчислените стойности със съответният „оптимален“ модел $P^*(M_j)$. Моделът с най-ниската стойност на MMR се дефинира като „оптимален“ модел съгласно този критерий, т.е.

$$\text{MMR} = \min_j (|R_j|), \quad (2.14)$$

където $R_j = \min_{a_j} \max_i |y_i - P^*_{(M_j)}(x_i, a_j)|$, е максималният остатък между изчислената и измерената стойност на зависимата променлива за $P^*(M_j)$, $j = 1, \dots, \kappa$, $i = 1, \dots, n$.

Хаусдорфовото разстояние като критерий за избор на модел. HD между две непразни множества $A = \{a_1, \dots, a_n\}$ и $B = \{b_1, \dots, b_m\}$ в метричното пространство се дефинира като:

$$\text{HD}(A, B) = \max(h(A, B), h(B, A)), \quad (2.15)$$

където:

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|, \quad (2.16)$$

и $\| \bullet \|$ е известна основна норма на A и B (например, Евклидова норма). Функцията $h(A, B)$ се нарича директно Хаусдорфово разстояние от A до B .

В дисертацията HD е използван като критерий наречен HDC за сравнение на множеството от точки на експерименталните данни A и множествата от точки $B^*(M_j), j = 1, \dots, k$, получени чрез изчисляване с индивидуалните „оптимални“ модели $P^*(M_j), j = 1, \dots, k$. Търсено е множество $B^* \in \{B_1, \dots, B_k\}$ за което $HDC(A, B^*) = \min HD(A, B^*(M_j)), j = 1, \dots, k$.

2.9. Подходи при разработване на софтуерните програми

Изборът на подходящи софтуерни програми със строго специфични цели е трудна задача. Такъв специфичен софтуер е именно програмата за изчисляване на критериите за избор на „оптимален“ модел и програмата за изчисляване на VO_{2max} . Това наложи разработване на специфична софтуерна програма за изчисляване на критериите за избор на „оптимален“ модел използвани в дисертацията, както и софтуерна програма за изчисляване на VO_{2max} с данни от Astrand-Ryhming тест. За реализирането на тези софтуерните програми е използван програмния език C#, в средата за разработка Visual C # 2010, с използването на библиотеката Framework 4.0.

Софтуерна програма „Comparing models“. Програмата „Comparing models“ изчисляваща критериите за избор на „оптимален“ модел (AIC, BIC, RMSD, MMR и HDC) използвани в дисертацията е реализирана като отделни модули, всеки от които служи за изчисляването на съответния критерий.

Софтуерна програма „VO2max Calculator “. Данните, с които потребителя използва тази програма оперира са информация за годините, мощността на натоварването и сърдечната честота в състояние на steady state, получени при субмаксимален тест на Astrand Rhythming за съответното изследвано лице. Програмата е разработена с възможност за изчисляването на VO_{2max} в (л/мин), както и в (мл/кг/мин), при допълнително въвеждане на информация за теглото на изследваното лице в килограми.

ТРЕТА ГЛАВА: РЕЗУЛТАТИ И ДИСКУСИЯ

В тази глава са представени резултатите от моделиране на зависимости при биомеханични изследвания, моделиране с регресионно уравнение за оценка на VO_{2max} и моделиране на лиганд-рецепторни взаимодействия, като е направено и обсъждане на получените резултати.

3.1. Моделиране на зависимости при биомеханични изследвания

Изометричните кривите на торг-ъглова позиция зависимости на екстензори, са моделирани с полиномни функции от различна степен и различни фитващи методи. В избора на модел са използвани информационни критерии AIC и BIC и RMSD критерий, както и предложените в дисертацията MMR и HDC. Изследвана е възможността за използване на MMR и HDC за избор на „оптимален“ модел от различни класове за фитване на торг-ъглова позиция зависимости на екстензорите на лакътя и е сравнена работата им с други често използвани критерии AIC, BIC и RMSD.

3.1.1. Резултати от статистически обработки на данни от измерване на зависимост торг-ъглова позиция в лакътна става

За да се трансформират получените експериментални данни за максималната волева контракция на екстензорите на лакътя в зависимост торг-ъглова позиция беше необходимо да се приложат следните изчислителни техники и статистически подходи: - нормализация на динамометрични данни чрез преизчисляване в проценти; - вариационен анализ за изчисляване на средни величини, стандартна грешка и вариационен коефициент. В Таблица 3.1. са представени стойностите на торга в абсолютни единици (Nm).

При нормализацията на стойностите за максималния изометричен торг на лицата при различни ъглови позиции, беше пресметнато процентното им участие спрямо максималната стойност на торга, независимо от ъгловата позиция в която е получена. Получените нормализирани стойности за торга в проценти са представени в Таблица 3.2.

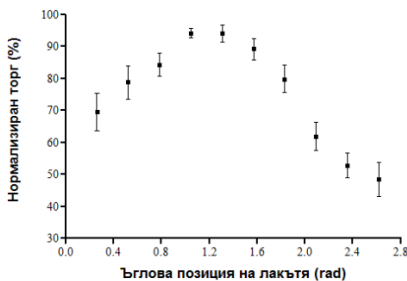
Таблица 3.1. Пиков торг на екстензорите в лакътя при здрави нетренирани жени, представен в абсолютни единици (Nm).

Изследвани лица	Ъглова позиция в лакътя (°)									
	15	30	45	60	75	90	105	120	135	150
01	27.60	30.10	28.70	30.00	32.10	26.60	28.80	22.10	22.90	23.80
02	17.40	21.30	23.10	24.20	24.50	25.50	19.00	15.30	11.00	7.90
03	15.30	18.70	24.20	29.10	34.50	29.50	20.40	14.80	17.40	11.40
04	20.90	21.00	19.30	30.20	29.00	28.50	24.30	20.40	14.90	14.10
05	14.50	16.40	24.60	26.90	28.80	29.20	23.90	19.00	16.20	13.80
06	28.60	34.60	35.60	36.80	36.80	37.30	37.50	33.80	28.40	27.80
07	36.10	37.20	33.90	34.10	29.50	31.10	27.70	21.30	21.10	23.80
08	13.00	13.70	12.20	15.60	12.20	14.50	15.60	11.00	5.90	5.70
09	12.90	19.70	24.60	28.00	30.30	20.20	20.60	14.60	12.20	9.70
10	26.70	29.00	33.80	32.20	32.00	29.30	23.70	16.50	16.20	15.50
$\bar{X}$	21.30	24.17	26.00	28.71	28.97	27.17	24.15	18.88	16.62	15.35
$\pm SEM$	2.54	2.53	2.29	1.85	2.15	1.96	1.94	1.99	2.02	2.35
V (%)	37.63	33.08	27.90	20.33	23.42	22.83	25.41	33.34	38.49	48.46

Таблица 3.2. Нормализирани данни (%) на пиков торг за екстензорите на лакътя при здрави нетренирани жени.

Изследвани лица	Ъглова позиция в лакътя (°)									
	15	30	45	60	75	90	105	120	135	150
01	85.90	93.70	89.40	93.40	100	82.80	89.70	68.80	71.30	74.10
02	68.20	83.50	90.50	94.90	96.00	100	74.50	60.00	43.10	30.90
03	44.30	54.20	70.10	84.30	100	85.50	59.10	42.80	50.40	33.00
04	69.20	69.50	63.90	100	96.00	94.30	80.40	67.50	49.30	46.60
05	49.60	56.10	84.20	92.10	98.6	100	81.80	65.00	55.40	47.20
06	76.20	92.20	94.90	98.10	98.10	99.40	100	90.10	75.70	74.10
07	97.00	100	91.10	91.60	79.30	83.60	74.40	57.20	56.70	63.90
08	83.30	87.80	78.20	100	78.20	92.90	100	70.50	37.80	36.50
09	42.50	65.00	81.10	92.40	100	66.60	67.90	48.10	40.20	32.00
10	78.90	85.70	100	95.20	94.60	86.60	70.10	48.80	47.90	45.80
$\bar{X}$	69.51	78.77	84.34	94.20	94.08	89.17	79.79	61.88	52.78	48.41
$\pm$SEM	5.88	5.16	3.55	1.48	2.63	3.30	4.27	4.36	3.96	5.29
V (%)	26.75	20.73	13.31	4.97	8.82	11.71	16.92	22.28	23.70	34.61

Осреднените данни и вариационните коефициенти за абсолютния и нормализирания торг при максимална изометрична контракция в различни ъглови позиции за екстензорите на лакътя при изследвания контингент са представени в Таблица 3.1. и Таблица 3.2. На Фигура 3.2 са показани графично средните стойности и стандартната грешка на нормализирания пиков торг за десетте участника в изследването, по отношение на десетте тествани позиции.



Фигура 3.2. Средни стойности на нормализирания торг ($\bar{X} \pm \text{SEM}$; %) на екстензорите в лакътя.

3.1.2. Резултати от моделиране с различни класове модели и критерии за оптималност. Адаптация на минимакс фитването в средата на Matlab

При фитване на експерименталните данни са използвани средните стойности на нормализираният пиков торг, осреднен за всички 10 участника, при всяка ъглова позиция (виж Таблица 3.3.).

Трябва да отбележим, че за тези фитвания измерените ъгли са изразени в радиани.

В настоящото изследване фитването на експериментални данни е направено с класове полиноми от втора, трета и четвърта степен (M_2 , M_3 и M_4):

$$M_2(x) = a + b * x + c * x^2;$$

$$M_3(x) = a + b * x + c * x^2 + d * x^3;$$

$$M_4(x) = a + b * x + c * x^2 + d * x^3 + e * x^4.$$

За всеки клас M_j , $j = 2, 3, 4$, е намерен „оптималният“ модел $P^*(M_j)$, с помощта на метода на най-малките квадрати, стабилна регресия и минимакс метод за фитване. „Оптималните“ модели P_2 , P_3 и P_4 от класовете M_2 , M_3 и M_4 , по отношение на метода на най-малките квадрати, стабилна регресия и минимакс метод, със стойностите на техните параметри и стойностите на RSS са показани в Таблица 3.4.

Таблица 3.3. Средни стойности на нормализирания торг ($\bar{X} \pm \text{SEM}$; %) на екстензорите на лакътя осреднени за всички 10 участници при всеки ъгъл.

	Ъгъл	Нормализиран торг [%]	
	rad	($\bar{X}$)	$\pm \text{SEM}$
15	0.2617994	69.51	5.88
30	0.5235988	78.77	5.16
45	0.7853982	84.34	3.55
60	1.0471980	94.20	1.48
75	1.3089970	94.08	2.63
90	1.5707960	89.17	3.30
105	1.8325960	79.79	4.27
120	2.0943950	61.88	4.36
135	2.3561940	52.78	3.96
150	2.6179940	48.41	5.30

Таблица 3.4. Стойностите на параметрите в различните модели за следните класове полиноми: M_2 - втора степен; M_3 - трета степен; M_4 - четвърта степен; Ъглите на лакътя са изразени в радиани за фитващите процедури.

Фитващ метод	Параметри					RSS	Полиномни модели
	a	b	c	d	e		
Най-малките квадрати	57.15	56.21	-23.80	-	-	197.00	M_2
	44.33	99.65	-59.78	8.33	-	128.10	M_3
	72.49	-38.25	133.00	-92.27	17.47	17.15	M_4
Стабилна регресия	52.49	68.64	-29.07	-	-	313.35	M_2
	45.14	97.96	-57.72	7.78	-	135.61	M_3
	71.34	-32.71	125.50	-88.44	16.81	17.45	M_4
Минимакс метод	66.29	41.92	-19.56	-	-	266.71	M_2
	42.51	101.50	-60.94	8.55	-	139.00	M_3
	76.36	-51.73	147.80	-98.67	18.39	21.58	M_4

При използваните методи за фитване и съответните софтуерни продукти, изчисляването на стандартни грешки и доверителни

интервали на параметрите е възможно единствено при метода на най-малките квадрати. Средните стойности на параметрите (със стандартни грешки и доверителни интервали) на „оптималните“ модели P_2 , P_3 и P_4 от класовете M_2 , M_3 и M_4 , по отношение на метода на най-малките квадрати са показани в Таблица 3.5.

Таблица 3.5. Средните стойности на параметрите на моделите (със стандартни грешки и доверителни интервали) по отношение на метода на най-малките квадрати за следните класове полиноми: M_2 - втора степен; M_3 - трета степен; M_4 - четвърта степен; Ъглите на лакътя са изразени в радиани за фитващите процедури.

Полиномни модели	Параметри	Средна стойност	Стандартна грешка	95% доверителен интервал
M_2	a	57.15	6.24	от 42.40 до 71.91
	b	56.21	9.96	от 32.66 до 79.75
	c	-23.80	3.37	от -31.76 до -15.83
M_3	a	44.33	8.97	от 22.39 до 66.27
	b	99.65	25.67	от 36.83 до 162.50
	c	-59.78	20.23	от -109.30 до -10.28
	d	8.33	4.63	от -3.00 до 19.67
M_4	a	72.49	6.12	от 56.76 до 88.22
	b	-38.25	26.35	от -106.00 до 29.48
	c	133.00	34.85	от 43.36 до 222.60
	d	-92.27	17.79	от -138.00 до -46.53
	e	17.47	3.07	от 9.57 до 25.36

Намирането на отделните „оптимални“ модели P_2 , P_3 и P_4 в класовете M_2 , M_3 и M_4 е направено с помощта на добре познат софтуер. Тук, в случая на използване на метод на най-малките квадрати и стабилна регресия, е използван GraphPad Prism. В случай на минимакс фитване е използван Matlab.

Адаптация на минимакс фитването в средата на Matlab

В Optimization Toolbox на Matlab има специална функция *fminimax*, която е използвана за изчисляване фитващите функции. Преди директното прилагане на функцията *fminimax* е дефинирана специална функция $F_j(x_i, a_j)$ във вида:

$$F_j(x_i, a_j) = |y_i - P_j(x_i, a_j)|, \quad (3.2)$$

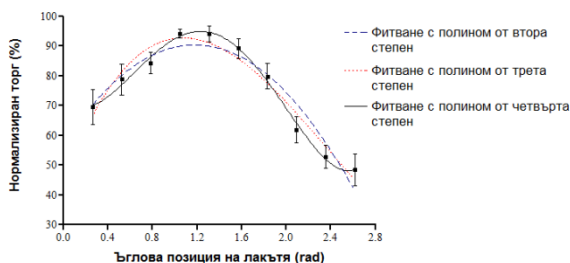
където

$$P_j(x_i, a_j) = a_{j0} + a_{j1} * x_i^1 + a_{j2} * x_i^2 + \dots + a_{jj} * x_i^j, \quad (3.3)$$

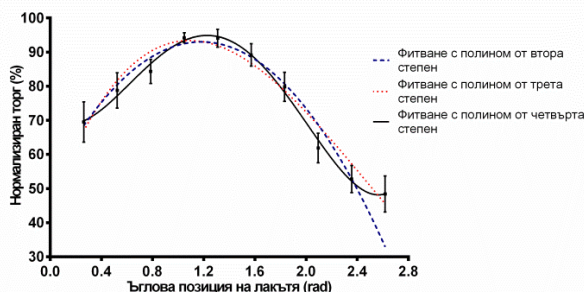
$y = (y_1, \dots, y_n)$ са значенията на зависимата променлива (максимален изометричен пиков торг), $x = (x_1, \dots, x_n)$ са значенията на независимата променлива (ъглова позиция), $a_j = (a_{j0}, a_{j1}, \dots, a_{jj})$ са параметрите в модела, $i = 1, \dots, n$, $j = 1, \dots, m$. В нашият случай $i = 1, \dots, 10$ и $j = 2, 3, 4$. След дефинирането на $F_j(x_i, a_j)$ е възможно прилагането на *fminimax* функцията за фитване на експерименталните

данни и са намерени оптималните модели P_2 , P_3 и P_4 на класовете M_2 , M_3 и M_4 , относно минимакс метод на фитване (виж Таблица 3.4.).

Фигура 3.3. и Фигура 3.4. представя графиките на „оптималните“ модели $P^*(M_j)$ на класовете M_j , $j = 2, 3, 4$, съответно по отношение на метода на най-малките квадрати и метода на стабилна регресия.



Фигура 3.3. Максимален изометричен торг-ъгъл криви за екстензорите на лакътя, получени при фитване с полиноми от втора, трета и четвърта степен, по метода на най-малките квадрати.



Фигура 3.4. Максимален изометричен торг-ъгъл криви за екстензорите на лакътя, получени при фитване с полиноми от втора, трета и четвърта степен, по метода на стабилна регресия.

За всеки от намерените „оптималните“ модели $P^*(M_j)$, $j = 2, 3, 4$, при всеки от фитващите методи, са изчислени често използваните в литературата критериите за избор на модел AIC, BIC и RMSD с помощта на разработената в дисертацията програма „Comparing models“ (Фигура 3.5.). Пресмятането на критерия AIC с помощта на тази програма, при сравняване на полиномите от трета и четвърта степен получени след метод на най-малките квадрати е показано на

Фигура 3.6. Стойностите на съответните критерии са дадени в Таблица 3.6., като най-добрите стойности са посочени в удебелен шрифт.

Съгласно критериите AIC, BIC и RMSD класът M_4 (полином от четвърта степен) съдържа „оптималния“ модел и при трите фитващи методи (виж Таблица 3.6.). Индивидуалните „оптимални“ модели $P^*(M_4)$, по отношение на различните фитващи методи са много близки по отношение на параметрите (виж Таблица 3.4.).

Таблица 3.6. Оценка на модели с критериите AIC, BIC и RMSD.

Фитващ метод	Полиномни модели			Критерии за избор на модел
	M_2	M_3	M_4	
Най-малките квадрати	45.81	50.50	45.39	AIC
	39.02	37.02	19.21	BIC
	5.31	4.62	1.85	RMSD
Стабилна регресия	50.45	51.07	45.57	AIC
	43.66	37.58	19.38	BIC
	6.69	4.75	1.87	RMSD
Минимакс метод	48.84	51.32	47.69	AIC
	42.05	37.83	21.51	BIC
	6.17	4.81	2.08	RMSD

В това изследване на зависимост торг-ъглова позиция, са въведени два нови критерия за избор на модел: MMR и HDC.

3.1.3. MMR като критерий за избор на оптимален модел от класове модели

За намерените индивидуалните „оптимални“ полиноми $P^*(M_j)$, $j = 2, 3, 4$, в избраните класове M_2 , M_3 и M_4 , за всеки от фитващите методи е минимизиран максималния остатък. Моделът с най-ниската стойност на MMR е избран за „оптимален“ модел съгласно този критерий.

3.1.4. Хаусдорфовото разстояние като критерий за избор на оптимален модел от класове модели

За изчисляване на Хаусдорфовото разстояние (HD) между множеството с експерименталните данни A и съответното множество от стойности B_j , изчислено с индивидуалните „оптимални“ модели $P^*(M_j)$, $j = 1, \dots, k$, за всеки от фитващите методи, в изследването е приложен следния алгоритъм представен като псевдокод:

input: Point Set A , Point Set B_j , $j = 1, \dots, k$, $i = 1, \dots, n$.

1. For any point a in A finding minimum distance to any point b_j in B_j ;
2. Finding the point a , in A with the largest minimum distance to any point in B ;
3. For any point b in B_j finding minimum distance to any point a_i in A ;
4. Finding the point b , in B_j with the largest minimum distance to any point in A ;
5. Finding the largest distance between directed Hausdorff distance $h(A, B_j)$ and directed Hausdorff distance $h(B_j, A)$.

output: HD between A and B_j .

Съгласно този алгоритъм и разработената програма „Comparing Models“ е изчислено HD между нашият еталон за сравнение (множеството от точки с експериментални данни) и всяко от множествата от точки получени чрез изчисляване с намерените индивидуални „оптимални“ полиноми $P^*(M_j)$, $j = 2,3,4$ и „оптималния“ модел, според HDC е намерен (виж Таблица 3.7). Моделът с най-ниската стойност на HD се дефинира като „оптимален“ модел съгласно този критерий. Резултатите от изчисленията на критериите MMR и HDC са дадени в Таблица 3.7. На Фигура 3.8. е показан диалоговият прозорец “Hausdorff distance between broken lines“ на програмата „Comparing Models“ при изчисляване на HD.

Таблица 3.7. Оценка на моделите с критериите MMR и HDC и сравнението им с информационните критерии AIC, BIC и RMSD, относно следните класове полиноми: M_2 - от втора степен; M_3 - трета степен; M_4 - четвърта степен.

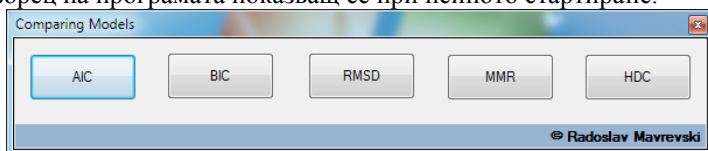
Фитващ метод	Полиномни модели			Критерии за избор на модел
	M_2	M_3	M_4	
Най-малките квадрати	45.81	50.50	45.39	AIC
	39.02	37.02	19.21	BIC
	5.31	4.62	1.85	RMSD
	8.60	5.45	2.36	MMR
	7.23	4.94	2.36	HDC
Стабилна регресия	50.45	51.07	45.57	AIC
	43.66	37.58	19.38	BIC
	6.69	4.75	1.87	RMSD
	15.46	6.71	2.40	MMR
	15.64	4.61	2.40	HDC
Минимакс метод	48.84	51.32	47.69	AIC
	42.05	37.83	21.51	BIC
	6.17	4.81	2.08	RMSD
	6.42	4.45	1.76	MMR
	6.42	4.45	1.76	HDC

Накрая са сравнени резултатите от предложените в дисертацията критерии за избор на модел MMR и HDC с резултатите от често използваните критерии AIC, BIC и RMSD за да се провери дали тези два нови критерия дават точни резултати и дали може да се използва като критерий за избор на модел (виж Таблица 3.7.). Най-добрите стойности за съответните критерии са посочени в удебелен шрифт.

3.1.5. Разработена софтуерна програма за пресмятане на критериите за избор на оптимален модел от класове модели

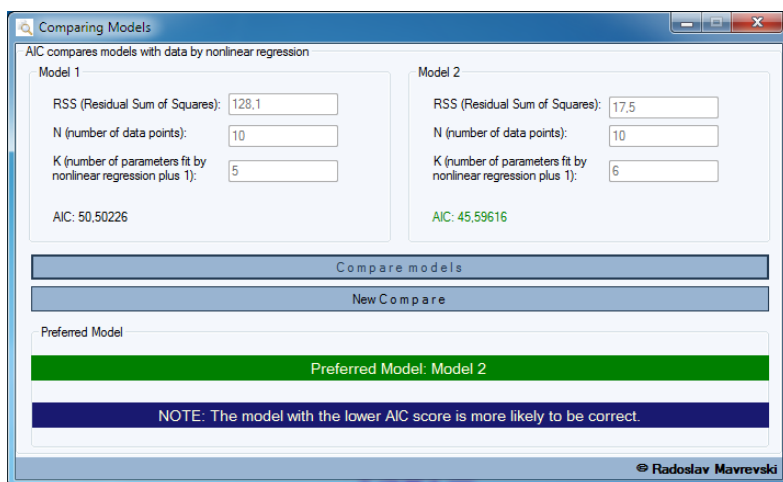
За изчисляване на стойностите на критериите AIC, BIC, RMSD, MMR и HDC за оценка и сравнение на „оптимални“ модели от различни класове, е използвана средата за разработка Visual C # 2010 и разработената в нея програма „Сравнение Models“ (Фигура 3.5.)

изчисляваща тези критериите. На Фигура 3.5. е показан главният прозорец на програмата показващ се при нейното стартиране.



Фигура 3.5. Основен прозорец на програмата „Comparing models“ за изчисляване на критериите за избор на модел.

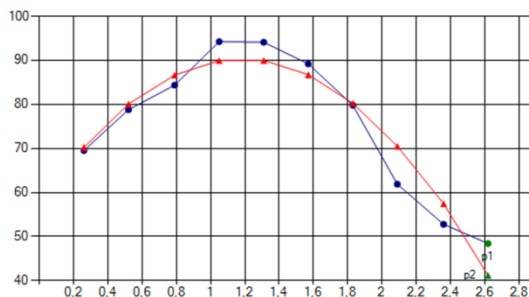
В диалоговият прозорец за изчисляване на AIC (Фигура 3.6.) за двата модела които се сравняват се въвеждат съответно остатъчната сума от квадратите (RSS), броя на точките от данни (N) и броя на параметрите в модела плюс един, тъй като регресия е “оценка” на сумата от квадратите, както и на стойностите на параметрите (K). Програмата позволява два модела да бъдат сравнявани или да се изчислят стойностите на няколко модела последователно и след това можем да изберем моделът за който стойността на AIC е най-ниска.



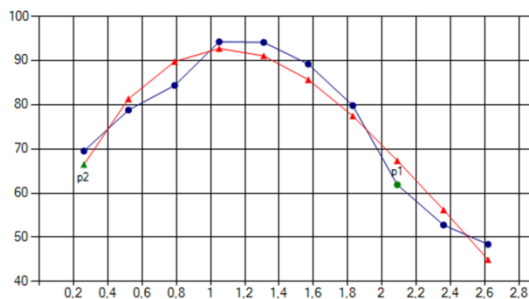
Фигура 3.6. Пример за изчисляване на AIC: диалогов прозорец на програмата „Comparing Models“ при изчисляване на AIC.

На Фигура 3.10. са представени графиките получени при изчисляване на $HD(A, B_j)$ за $j = 2, 3, 4$ в диалоговият прозорец „Hausdorff distance between broken lines“ от програмата „Comparing Models“.

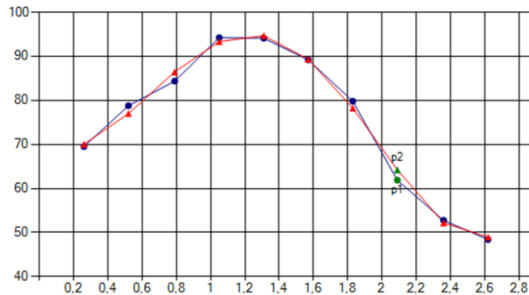
а)



б)



в)



Фигура 3.10. Графики, създадени в диалоговият прозорец „Hausdorff distance between polylines“ от програмата „Comparing Models“ при изчисляване на Хаусдорфовото разстояние (HD) между множеството от точки на експериментални данни - A (кръгови точки, синя линия) и множествата от точки изчислени с полиномите (триъгълни точки, червена линия) от втора (а; $HD = 7.23$), трета (б; $HD = 4.94$) и четвърта (в; $HD = 2.36$) степен, при използване на фитване по метода на най-малките квадрати; точките от двете графики, където се постигат разстоянието са означени като p_1 и p_2 (зелените точки).

3.1.6. Дискусия

При това изследване е установено, че: (1) „оптималният“ модел описващ зависимостта торг-ъглова позиция на екстензорите на лакътя е полином от четвърта степен; (2) MMR и Хаусдорфовото разстояние може да се използват като критерий за оценка на „оптимален“ модел, в сравнение с AIC, BIC и RMSD; и (3) едно от предимствата на критериите MMR и HDC над AIC, BIC и RMSD е, че те не използват остатъчната сума от квадрати, която не винаги е на разположение директно от софтуерните пакети и може да се използват при избора на модел след всеки фитиващ метод с еднаква лекота.

В изследванията си (Pinter I., B. Maarten, S. Jeroen, 2010) установят, че „оптимален“ клас модел за торг-ъглова позиция на лакътя, за екстензори и флексори е трета степен полином. Това може би се дължи на факта, че изследваните лица са от смесен тип (мъже и жени), докато в нашите изследвания участваха само жени. Освен това в определянето на „оптималния“ клас модел, те използват само критерия AIC.

В случай на най-малките квадрати, препоръчително е сравняване на моделите с AIC и BIC, защото остатъчната сума от квадратите е достъпно директно и не изискват допълнителни изчисления, както е случаят с другите фитиващи методи.

Очевидно е, че MMR е по-подходящ критерий за сравняване на модели след минимакс фитване.

Методът на RMSD (Myung I. J., 2000) не би могъл да се справи при случаи с малки размери на пробите и голям брой на параметри в моделите, тъй като това може да доведе до отрицателни квадратни корени. Освен това, няма статистическа обосновка за RMSD.

Задоволителен компромис между доброта на фитване и пестеливост може да се постигне чрез прилагане на AIC, BIC и RMSD. Те са чувствителни само към един от аспектите на сложност, брой параметри, но не са чувствителни към функционална форма на модела. За това е необходимо оценяването на моделите да се прави като се комбинират различни критерии за оценка на „оптималност“ на моделите.

HD беше успешно приложено като критерий HDC за избор на „оптимален“ модел. Числените резултати, получени в това изследване са окуражаващи, така че вероятно HDC може да се използва като критерий и в други медико-биологични изследвания. MMR и HDC, също както AIC, BIC и RMSD, избират клас M_4 с голяма разлика в оценките в сравнение с M_2 и M_3 .

3.2. Резултати от моделиране с регресионно уравнение за определяне на VO_{2max}

Разликата между стойностите на VO_{2max} , определени с Astrand-Ryhming тест и с максималния спирогеометричен тест в направените изследвания, наложи модифициране на модела за пресмятане, чрез прецизиране на параметрите в регресионното уравнение, както и изработване на съответен софтуер улесняващ пресмятането на VO_{2max} с това уравнение. Модифицирайки параметрите на модела за нашият контингент, беше получена средна стойност за $VO_{2max} = 3.326 \pm 0.480$ л/мин, която не се различава статистически ($p=0.838$) от тази при максималния тест $VO_{2max} = 3.280 \pm 0.337$ л/мин.

3.2.1. Резултати от статистически обработки на данни от тестове за определяне на VO_{2max}

Осреднените данни със стандартно отклонение и вариационните коефициенти за VO_{2max} измерена чрез максималния тест при изследвания контингент са представени в Таблица 3.8.

Таблица 3.8. Експериментални данни ($\bar{X} \pm SD$; $V\%$) за VO_{2max} измерена чрез максимален тест

Изследвано лице	Години	VO_{2max} измерена (л/мин)
01	23	2.999
02	20	3.435
03	20	2.687
04	22	4.090
05	20	3.223
06	19	2.755
07	19	2.591
08	21	2.779
09	24	3.018
10	19	4.090
11	20	2.938
12	20	3.716
13	20	4.016
14	24	3.680
15	19	3.826
16	19	3.348
17	25	3.410
18	25	3.175
19	20	3.412
$\bar{X}$		3.326
$\pm SD$		0.480
$V\%$		14.450

Осреднените данни със стандартно отклонение и вариационните коефициенти за VO_{2max} изчислена чрез Astrand-Ryhming тест и регресионно уравнение при изследвания контингент са представени в Таблица 3.9.

Таблица 3.9. Експериментални данни ($\bar{X} \pm SD$; $V\%$) за VO_{2max} изчислена чрез Astrand-Ryhming тест.

ИЛ	Години	Мощност на натоварването (W)	HRsm (уд/мин)	VO_{2max} изчислена (л/мин)
01	23	150	185	2.303
02	20	170	156	3.424
03	20	145	170	2.601
04	22	200	177	3.189
05	20	170	169	3.012
06	19	155	180	2.541
07	19	180	175	3.021
08	21	165	172	2.835
09	24	165	178	2.631
10	19	135	184	2.185
11	20	150	172	2.630
12	20	190	186	2.869
13	20	180	157	3.562
14	24	195	169	3.300
15	19	185	171	3.207
16	19	180	168	3.219
17	25	200	182	2.990
18	25	150	172	2.535
19	20	165	170	2.908
$\bar{X}$		170.00	173.30	2.893
$\pm SD$		19.29	8.28	0.376
$V\%$		11.35	4.78	13.030

Съкращения: ИЛ – изследвано лице; HRss – сърдечна честота при steady state

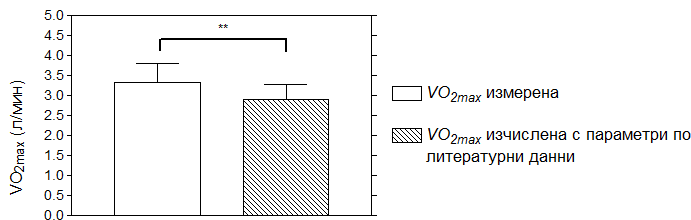
3.2.2. Резултати от моделиране с регресионно уравнение за определяне на VO_{2max} с тест на Astrand-Ryhming

Стойностите на VO_{2max} на база на Astrand-Ryhming тест (виж Таблица 3.9.) са изчислявани с регресионно уравнение за мъже по (Adams G., 1998), което се използва широко в практиката. Уравнението е следното:

$$VO_{2max} = ((P * a) + b) * \frac{(220 - age) - 61}{HRsm - 61}, \quad (3.4)$$

където VO_{2max} е максималната кислородна консумация (л/мин), P е мощността на натоварването измерена във ватове (Bm), age е възрастта на изследваните лица (години), $HRss$ е сърдечната честота (heart rate) в състояние на steady state (уд/мин) и параметрите a и b по литературни данни са съответно $a=0.012$ и $b=0.3$.

За да проверим различията между измерените и изчислените стойности на VO_{2max} с двата теста (максимален тест и Astrand-Ryhming) са сравнени техните средни стойности с непараметричен тест на Mann-Whitney за сравняване на средни при малки извадки. Установени са статистически достоверни различия (Фигура 3.11.).



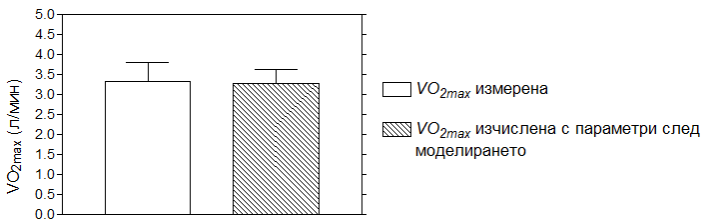
Фигура 3.11. Оценка на различията в стойностите на VO_{2max} между максимален тест до отказ (измерени стойности) и Astrand-Ryhming тест (изчислени стойности); ** статистически достоверна разлика ($p < 0.05$) с Mann Whitney тест преди моделирането.

За определяне на VO_{2max} с теста на Astrand-Ryhming за определения контингент изследвани лица, бе направено математическо моделиране на уравнението с множествен регресионен анализ по метода на най-малките квадрати със софтуерен пакет SPSS Statistics. След направеният регресионен анализ са получени нови стойности на параметрите в уравнението за изчисляване на VO_{2max} с тест на Astrand-Ryhming за изследваният контингент. Новите стойности на параметрите в модела са дадени в Таблица 3.10.

Таблица 3.10. Стойности на параметрите в уравнението за пресмятане на VO_{2max} при субмаксимален тест на Astrand Rhythmic при мъже

Параметър	По литературни данни	Нови стойности след моделиране ($\bar{X} \pm SEM$)
a	0.012	0.008 ± 0.005
b	0.300	1.294 ± 0.890

Оценката на средните стойности за VO_{2max} поучени с максимален тест и Astrand Rhythmic тест с параметри a и b след моделирането показва, че не се наблюдават статистически достоверни различия между тях (Фигура 3.13).



Фигура 3.13. Оценка на различията в стойностите на VO_{2max} между максимален тест до отказ (измерени стойности) и Astrand-Ryhming тест (изчислени стойности); не се установява статистически достоверна разлика с Mann Whitney тест след моделирането.

3.2.3. Разработена софтуерна програма за пресмятане на VO_{2max} с тест на Astrand-Ryhming

За изчисляване на стойностите на VO_{2max} при субмаксимален тест на Astrand Rhythmic при мъже е използвана средата за разработка Visual C # 2010 и разработената в нея програма „VO2max Calculator“ (Фигура 3.14.) изчисляваща VO_{2max} .

Фигура 3.14. Софтуерна програма за изчисляване на VO_{2max} при субмаксимален тест на Astrand Rhythmic при мъже.

3.2.4. Дискусия

Данните от това моделиране показват, че при редица измервания на показатели за ниво на тренираност, здравен статус и други, при функционални изследвания в практиката се избягват сложните и трудоемки методологични подходи и се заместват с регресионни уравнения за предсказването на тези показатели. В тези случаи винаги се залагат някакви експериментално установени данни и когато контингентът се променя и регресионните уравнения стават по-малко точни и е необходимо тяхното прецизиране. Типичен пример е настоящото изследване което показва, че за изследвания контингент е необходимо параметрите в регресионното уравнение периодично да бъдат проверявани. Ето защо в случая те бяха прецизирани за съответния контингент изследвани нетренирани лица. Нещо повече, беше създадена софтуерна програма, с която лесно да бъдат изчислени стойности на VO_{2max} чрез моделираното уравнение и резултати от Astrand Rhythmic тест.

3.3. Резултати от моделиране на зависимости между *in silico* и *in vitro* изследвания на лиганд-рецепторни взаимодействия

Особено важно за намиране на селективни и ефективни лиганди за даден рецептор с използването на компютърни методи е да се подбере подходящата програма за докинг и оценъчната функция, която би могла да определи точния начин на свързване. Това може да бъде

оценено като се намери връзка между резултатите от докинга и биологичните експерименти. В настоящото изследване е направен опит да се намери корелация между данните получени от докинга и *in vitro* тестовите, като тази корелация би могла да се използва по-нататък за дизайн на нови аналози на енкефалините с по-силен биологичен ефект. Обектите, използвани в нашата работа са енкефалинови аналози (Таблица 3.12.) и теоретичен модел на ДОР (PDBe id: 1ozc).

Таблица 3.12. Използвани лиганди при моделиране на лиганд-рецепторно действие с помощта на докинг изследвания

№	Лиганд	Аминокиселинна последователност
1	DPDPE	Tyr-D-Pen-Gly-Phe-D-Pen
2	[Leu ⁵]-enk	Tyr-Gly-Gly-Phe-Leu
3	[Met ⁵]-enk	Tyr-Gly-Gly-Phe-Met
4	[Cys(Bzl) ² , Leu ⁵]-enk	Tyr-Cys(Bzl)-Gly-Phe-Leu
5	[Cys(Bzl) ² , Met ⁵]-enk	Tyr-Cys(Bzl)-Gly-Phe-Met
6	[Cys(O ₂ NH ₂) ² , Leu ⁵]-enk	Tyr-Cys(O ₂ NH ₂)-Gly-Phe-Leu
7	[Cys(O ₂ NH ₂) ² , Met ⁵]-enk	Tyr-Cys(O ₂ NH ₂)-Gly-Phe-Met
8	[DCys(O ₂ NH ₂) ² , Leu ⁵]-enk	Tyr-D-Cys(O ₂ NH ₂)-Gly-Phe-Leu
9	[DCys(O ₂ NH ₂) ² , Met ⁵]-enk	Tyr-D-Cys(O ₂ NH ₂)-Gly-Phe-Met
10	[HCys(O ₂ NH ₂) ² , Leu ⁵]-enk	Tyr-HCys(O ₂ NH ₂)-Gly-Phe-Leu
11	[HCys(O ₂ NH ₂) ² , Met ⁵]-enk	Tyr-HCys(O ₂ NH ₂)-Gly-Phe-Met

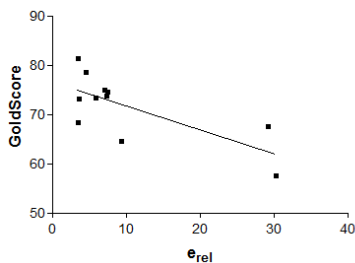
Проведен е докинг на енкефалиновите аналози с ДОР с използването на две програми за докинг (MVD и GOLD).

В резултат от *in vitro* тестовите са получени стойностите за IC₅₀, K_A и e_{rel}. Тези данни са представени в Таблица 3.13. Направен е опит за намиране на корелация между данните от докинга и *in vitro* експериментите. За да се намери връзка между тези параметри се търси корелация между тях.

Таблица 3.13. Стойности на тоталната енергия, фитнес функцията, IC₅₀, K_A и e_{rel} на изследваните лиганди

Лиганд	E-Total (кДж/мол)	GoldScore	IC ₅₀ (нмол/л)	K _A (нмол/л)	e _{rel}
[Cys(Bzl) ² , Leu ⁵]-enk	-71.23	64.68	8.30	68.50	9.30
[Cys(Bzl) ² , Met ⁵]-enk	-90.93	81.49	9.53	23.80	3.50
[Cys(O ₂ NH ₂) ² , Leu ⁵]-enk	-111.60	67.72	1.29	36.40	29.20
[Cys(O ₂ NH ₂) ² , Met ⁵]-enk	-41.39	73.91	2.22	14.10	7.30
[DCys(O ₂ NH ₂) ² , Leu ⁵]-enk	-110.93	74.73	11.40	73.40	7.40
[DCys(O ₂ NH ₂) ² , Met ⁵]-enk	-109.22	75.13	75.96	463.00	7.10
DPDPE	-127.27	57.67	6.18	180.00	30.20
[HCys(O ₂ NH ₂) ² , Leu ⁵]-enk	-59.66	68.43	31.92	76.40	3.40
[HCys(O ₂ NH ₂) ² , Met ⁵]-enk	-44.79	78.65	16.09	55.70	4.50
[Leu ⁵]-enk	-137.50	73.42	11.45	54.90	5.80
[Met ⁵]-enk	-107.05	73.26	18.91	48.40	3.60

Използвайки данните от Таблица 3.13 е направен опит да се намери корелация между стойностите, получени от докинга и от биологичните изследвания. В случая, когато се използва MVD и се получават стойности за тоталната енергия, най-високи корелационни коефициенти се получават с e_{rel} , съответно коефициент на Спирмън (-0.43) и на Пирсън (-0.41). В този случай корелацията е умерена до ниска и не би могла да се използва категорично за определяне на връзка между тези данни. В случая на GOLD се получават данни за фитнес функцията. Корелационните коефициенти са съответно Спирмън (-0.53) и Пирсън (-0.72). В този случай корелационния коефициент на Пирсън показва силна зависимост между двете множества от данни. Особено важно е да се намери не каквато и да е връзка между тези данни, а връзката между тях трябва да е линейна. Тогава може да се каже, че ако едната стойност расте, то и другата ще нараства (при положителен корелационен коефициент) и в обратния случай, когато едната стойност расте другата ще намалява (при отрицателен корелационен коефициент). Ето защо значима корелация е тази между фитнес функцията и e_{rel} (Фигура 3.17.). В дадения случай с нарастване на фитнес функцията ефикасността намалява. От гледна точка на фармакологичната активност на лиганда, тази зависимост може да се обясни по следния начин. Ако един лиганд се свързва здраво с рецептора, той блокира дадения рецептор за по-продължителен период от време. Това от своя страна прави концентрацията на свободните рецептори по-ниска, т.е. няма налични свободни рецептори, които да се свързват с постъпващите лиганди. Един лиганд е по-ефективен, ако за кратко успее да се свърже с рецептора, предизвиква очаквания ефект и бързо след това успее да напусне свързващия център на рецептора, освобождавайки място за следващия лиганд.



Фигура 3.17. Корелация между стойностите на фитнес функцията *GoldScore* от GOLD и e_{rel} на изследваните лиганди.

ИЗВОДИ

1. Предложена е една последователна стратегия при анализа и моделирането на експериментални данни при медико-биологични изследвания;

2. Предложен е „оптимален“ математически модел на торг-ъглова позиция зависимости в областта на биомеханиката, базирайки се на различни фитващи методи като най-малките квадрати, стабилна регресия и минимкас фитване, както и подходящ софтуер за регресионен анализ;

3. Намереният „оптимален“ модел подпомага биомеханичния анализ на поведението на екстензорите на лакътя за генериране на максимална изометрична сила съобразно изходната дължина на мускулите;

4. Приложени са известните в литературата критерии AIC, BIC и RMSD за избор на „оптимален“ модел от различни класове модели описващ торг-ъглова позиция зависимости;

5. Предложени са два нови критерия MMR и HDC и е оценена тяхната ефективност чрез сравняването им с AIC, BIC и RMSD;

6. Разработена е софтуерна програма за пресмятане на критериите AIC, BIC, RMSD, MMR и HDC за избор на „оптимален“ модел от различни класове модели;

7. В резултат от направеният многофакторен регресионен анализ с SPSS за оценка на параметрите в регресионното уравнение за пресмятане на VO_{2max} са намерени „оптималните“ параметри в модела при нетренирани лица;

8. Разработена е софтуерна програма за пресмятане на VO_{2max} с тест на Astrand-Ryhming, която улеснява определянето на VO_{2max} с този индиректен тест;

9. Установена е голяма корелация ($r = -0.72$) между данните получени от докинг изследвания с GOLD и *in vitro* измервания; това доказва надеждността на резултатите при използване на GOLD и ДОР при сравнително изследване с MVD.

СПИСЪК С ПУБЛИКАЦИИ ПО ДИСЕРТАЦИЯТА

Публикации в списания:

1. Dzimbova T., R. Mavrevski, N. Pencheva, T. Pajpanova, P. Milanov. Computer modeling of ligand-receptor interaction – encephalin analogues delta-opioid receptor. *Bulgarian chemical communications*, 2012, 44(3), 242-246 (Impact factor: 0.320).
2. Mavrevski R., Selection and comparison of regression models: estimation of torque-angle relationships. *C. R. Acad. Bulg. Sci.*, 2014, 67(10), 1345-1354. (Impact factor: 0.198).
3. Mavrevski R., G. Koroleova, N. Pencheva, P. Milanov, B. Yurukov. The Hausdorff distance as a criterion for the optimal model selection in torque-angle relationships. *Acta of bioengineering and biomechanics*. (In Preparation).
4. Манчев А., Р. Мавревски, Н. Пенчева, П. Миланов, Ч. Коцев Сравнителна оценка на максимален тест до отказ и Astrand-Rhiming тест за определяне на VO_{2max} . *Спорт и наука*, 2012, Том 5, 140-148.

Публикации в сборници на конференции:

5. Mavrevski R. Statistical modelling in food technology. Proceedings of the Scientific Conference “Enhancement of the competitiveness and restructuring of the food subsectors through benchmarking”, 15 May 2014, Blagoevgrad, Bulgaria, Volume, pp. 118-122.
6. Mavrevski R., A. Stoilov. Weather station at SWU – statistical analysis and reports. Proceedings of the Fourth International Scientific Conference – FMNS2011, 2011, Blagoevgrad, Bulgaria, Volume 1, pp. 564 -569.
7. Metodi Traykov, Radoslav Mavrevski, Anton Stoilov. Pricing of goods and services through partial differential equations. Proceedings of the Scientific Conference “Enhancement of the competitiveness and restructuring of the food subsectors through benchmarking”, 15 May 2014, Blagoevgrad, Bulgaria, Volume, pp. 74-83.
8. Stoyanov K., A. Stoilov, A. Pulev, L. Sakelarieva, R. Mavrevski. Temperature conditions of the soils in the Blagoevgradska Bistritsa River basin (Rila mountain) reports. Proceedings of the Fourth International Scientific Conference – FMNS2011, 2011, Blagoevgrad, Bulgaria, Volume 1, pp. 583 -589.
9. Stoilov A., R. Mavrevski. A benchmark for docking of 60 protein ligand complexes. 14th National Biostatistics Congress with International Participation, 4-7 September 2012, Kadir Has Center, Kayseri, Turkey. Poster presentation.

УЧАСТИЕ В ИНТЕНЗИВНИ КУРСОВЕ И УЧАСТИЕ В ПРОЕКТИ

Участие в интензивни курсове

- DAAD Intensive course "Symmetry in Science and Art", 10-16 May, 2011, Vrnjacka Banja, Serbia;
- DAAD Intensive course "Applications of the Calculus of Variations and Optimal Control. The Smooth and Nonsmooth Cases", 11-18 July, 2011, Struga, Macedonia;
- DAAD Intensive course "Robotics and Mathematics", 12-18 August, 2012, Ohrid, Macedonia;
- DAAD Intensive course "Summability Theory and Statistical Convergences", 20-27 August, 2012, Pristina, Kosovo;
- Курс Суперкомпютърни приложения в природните науки, „Използване на системната и приложната среда на Blue. Моделиране на взаимодействия на биомолекули“, 18-19 февруари, Български суперкомпютърен център, 2012 София, България;
- Training School "New Drugs for Neglected Diseases", 15-20 October, 2012 Siena, Italy.

Участие в проекти

- “Биоинформатични изследвания върху свойствата и биологичната активност на протени и лекарствено-рецепторни взаимодействия”, финансиран по договор ДО 02-162/16.12.2008 г. към ФНИ на РБългария;
- „Фармакологични и биоинформатични изследвания върху съединения с антипаразитна активност”, по програма COST ДО 02-135 (август 2009);
- “Enhancement of the competitiveness and restructuring of the food subsectors through benchmarking” (F.IND. CONSULTING). Договор B2.32.01/17.02.2012 по Европейската програма за териториално сътрудничество "Гърция-България 2007-2013 г.";
- “Interregional Cooperation at Scientific Computing in Interdisciplinary Science (ICoSCIS)”. Договор B2.33.02/28.02.2012 по Европейската програма за териториално сътрудничество "Гърция-България 2007-2013 г.";
- „Оползотворяване на слънчевата енергия в Югозападна България”, финансиран по договор ДМУ 03/1 (март 2012) към ФНИ на РБългария.
- „Биоинформатични изследвания: нагъване на протеини, докинг и предсказване на биологична активност“, финансиран по договор I02/16, 12.12.14 към ФНИ на РБългария.

ЛИТЕРАТУРА

Acquah de-Graft H., 2010. Comparison of Akaike information criterion (AIC) and Bayesian information criterion (BIC) in selection of an asymmetric price relationship. *Journal of Development and Agricultural Economics*, 2(1), pp. 001-006.

Adams G., 1998. Aerobic cycling. In: *Exercise Physiology Laboratory Manual*. New York: McGraw-Hill, pp. 139-153.

Akaike H., 1973. *Information theory as an extension of the maximum likelihood principle*, p. 267–281 in B. N. Petrov, and F. Csaki, (eds.) *Second International Symposium on Information Theory*, Budapest: Akademiai Kiado.

Akaike H., 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, Volume 19, pp. 716-723.

Armstrong J. S., F. Collopy, 1992. Error Measures For Generalizing About Forecasting Methods: Empirical Comparisons. *International Journal of Forecasting*, 8(1), pp. 69-80.

Astrand P-O, 2002. Измерване на издръжливостта. Фактори за измерване. От: *Издръжливостта в спорта*. София: Тип-Топ Принт, pp. 233-235.

Barlow R. B., S. M. Bond, C. Grant, D. S. McQueen, Z. Yaqoob, 2001. A comparison of effects measured with isotonic and isometric recording: II. Concentration-effect curves for agonist. *British Journal of Pharmacology*, Volume 133, pp. 1087-1095.

Barlow R.B., S.M. Bond, E. Bream, L. Macfarlane, D.S. McQueen, 1997. Antagonist inhibition curves and the measurement of dissociation constants. *British Journal of Pharmacology*, Volume 120, pp. 13-18.

Barlow R.B., 1975. Use of the logistic function for the calculation of dose-ratios and potency ratios. *British Journal of Pharmacology*, Volume 53, pp. 139-141.

Bober T., Kulig K., Burnfield J., Pietraszewski B., 2002. Predictive torque equations for joints of the extremities. *Acta. Bioeng. Biomech.*, Volume 4, pp. 49-61.

Box G.E.P., W.G. Hunter, J.S. Hunter, 1978. *Statistics for experimenters*. New York: JohnWiley and Sons.

Brughelli M., J. Cronin, K. Nosaka, 2010. Muscle Architecture and Optimum Angle of the Knee Flexors and Extensors: A Comparison Between Cyclists and Australian Rules Football Players. *Journal of Strength and Conditioning Research*, 24(3), pp. 717-721.

Burnham K. P., D. R. Anderson, 2002. *Model Selection and Multimodel Inference*. 2 ed. New York: Springer-Verlag.

Daniel P., A. Gregory, J. William, 1993. Comparing Images Using the Hausdorff Distance. *IEEE Transactions on pattern analysis and machine intelligence*, 15(9), pp. 850-863.

Ekins S., J. Mestres, B. Testa, 2007. In silico pharmacology of drug discovery: methods of virtual ligand screening and profiling. *Br J Pharmacol*, 152(1), pp. 9-20.

Enoka R.M., 2002. Muscle Mechanics. In: *Neuromechanics of Human Movement*. In: 3 ed. Champaign, IL: Human Kinetics, pp. 257-278.

Hyndman R. J., A. B. Koehler, 2006. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), pp. 679-688.

Joseph B., L. Nicole, 2004. Methods and Criteria for Model Selection. *Journal of the American Statistical Association*, 99(465), pp. 279-290.

Konishi S., G. Kitagawa, 2008. *Information criteria and statistical modeling*. New York: Springer Science and Business media.

Kullback S., R. A. Leibler, 1951. On information and sufficiency. *Annals of Mathematical Statistics*, Volume 22, pp. 79-86.

Lunneborg C.E., 1994. *Modeling experimental and observational data*. Belmont, CA: Duxbury Press.

Manocha D., C. Erikson, 1998. GAPS: General and automatic polygonal simplification. In *Proc. of ACM Symposium on Interactive 3D Graphics*, pp. 79-88.

Mason R. L., R. F. Gunst, J. L. Hess, 2003. *Statistical Design and Analysis of Experiments*. New Jersey: Published by John Wiley & Sons, Inc., Hoboken.

Motulsky H., R. Brown, 2006. Detecting outliers when fitting data with nonlinear regression – a new method based on robust nonlinear regression and the false discovery rate. *BMC Bioinformatics*, Volume 7, pp. 123-143.

Myung I. J., 2000. The importance of complexity in model selection. *Journal of Mathematical Psychology*, Volume 44, pp. 190-204.

Nutanong S., E. H. Jacox, H. Samet, 2011. *An Incremental Hausdorff Distance Calculation Algorithm*. In: *Proceedings of the VLDB Endowment*, 4(8), pp. 506-517.

Pencheva N., M. Kokova, S. Dencheva, 2013. *Torque – angle relationships of the elbow flexors and extensors in healthy females*. Barcelona, Book of Abstracts of 18th annual Congress of the European College Of Sport Science, 26th - 29th June.

Pencheva N., P. Milanov, L. Vezencov, T. Pajpanova, E. Naydenova., 2004. Opioid profiles of Cys2-containing enkephalin analogues. *Eur. J. Pharmacol.*, Tom 498, pp. 249-256.

Pinter I., B. Maarten, S. Jeroen, 2010. Isometric torque–angle relationships of the elbow flexors and extensors in the transverse plane. *Journal of Electromyography and Kinesiology*, 20(5), pp. 923-931.

Powell M. J. D., 1981. *Approximation Theory and Methods*. London: Cambridge University Press.

Rabinowitz P., 1968. Applications of linear programming to numerical analysis. *SIAM Rev*, Volume 10, pp. 121-159.

Sadanori K., G. Kitagawa, 2008. *Information Criteria and Statistical Modeling*. New York: Springer Science+Business Media, LLC.

Saltin B, S.O. Strange, 1991. Maximal oxygen uptake: “old” and “new” arguments for a cardiovascular limitation. *Med Sci Sports Exerc.*, Volume 24, pp. 30-37.

Schwarz G., 1978. Estimating the dimension of a model. *Annals of Statistics*, Volume 6, pp. 461-464.

Shapiro M., M. Blaschko, 2004. On Hausdorff Distance Measures. University of Massachusetts Amherst, Computer Science Dept. Technical report.

Shenk T.M., A.B. Franklin, 2001. *Modeling in natural resource management*. Washington: Island Press.

Siconolfi S.F, C.E. Garber, T.M. Lasater, R.A. Carleton, 1985. A simple, valid step test for estimating maximal oxygen uptake in epidemiologic studies. *Am J Epidemiol.*, Volume 121, pp. 382-390.

Thomsen R., M. H. Christensen, 2006. MolDock: a new technique for high accuracy molecular docking. *Journal of Medicinal Chemistry*, Volume 49, pp. 3315-3321.

Van den Bos A., 1988. Nonlinear least-absolute-values and minimax model fitting. *Automatica*, 24(6), pp. 803-808.

Wadood A., N. Ahmed, L.Shah, A. Ahmad, H. Hassan, S. Shams, 2013. In-silico drug design: An approach which revolutionarised the drug discovery process. *OA Drug Design & Delivery*, 1(1), pp. 1-4.

Weisberg S., 1985. *Applied linear regression*. 2 ed. New York: Wiley.

Wilson E., E. Rustighi, P. Newland, B. Mace, 2012. A comparison of models of the isometric force of locust skeletal muscle in response to pulse train inputs. *Biomech. Model Mechan.*, Volume 11, pp. 519-532.

Zatsiorsky V.M., Prilutsky B.I., 2012. *Biomechanics of skeletal muscles*. Champaign, IL: Human Kinetics.

Zucchini W., 2000. An Introduction to Model Selection. *Journal of Mathematical Psychology*, Volume 44, pp. 41-61.

Душков В., Стефанова Д., Джарова Т., 1986. *Функционални изследвания в спорта и масовата физкултура*. София: МФ.

Манчев А., 2014. *Параметри на аеробния капацитет и нормиране на натоварването за бегова издръжливост на студенти*. Благоевград: Дисертация.

Пенчева Н., 2010а. *Неврофизиология на движенията, Магистърска програма по Кинезиология: оп line помагало за обучение*. Благоевград: ЮЗУ "Неофит Рилски".

Пенчева Н., Коцев Ч., Манчев А., 2010. Роля на тестове за аеробен капацитет за обучението по спорт при здрави нетренирани студенти. „Съвременното образование – мисия и визии“, *Сборник научни студии и статии*, Университетско издателство „Неофит Рилски“, Благоевград, pp. 469-475.

Пенчева Н., 2010б. *Спироергометрия, изокинетична динамометрия и кинезиологична електромиография, Магистърска програма по кинезиология и спортна кинезиология: оп line помагало за обучение*. Благоевград: ЮЗУ "Неофит Рилски".

Попов Н., 2003. *Клинична патокинезиологична диагностика в ортопедично-травматологичната кинезитерapia*. София: НСА-ПРЕС.

Флойд Р. Т., 2008. *Наръчник по анатомична кинезиология*. София: Медицина и физкултура.

ANNOTATION IN ENGLISH

Medico-biological research are related to the design of biological experiments, as well as the collection, compilation and analysis of data from these experiments and the interpretation of the results and the conclusions we can draw from them. All three, statistical methods, mathematical, and computer modeling are crucial for the medico-biological research. An important challenge for the mathematical modeling involves determining the appropriate class model that should be used to draw conclusions and predictions.

The need for mathematical and computer modeling, in medico-biological research, is based on the fact that the immediate examination of a large number of objects is either impossible or requires a lot of time, money, and specialized technical equipment for experiments. In practice, the choice of a model and the additional assessment are major issues. A key aspect is the choice of the „optimal“ model from a set of candidate class models.

The studies were conducted at the Center for functional studies in the sport and kinesitherapy at the Faculty of Public Health and Sports of the South-West University „Neofit Rilski“ in Blagoevgrad, where such experiments are constantly performed, and thus require such modeling and optimization.

The focus of the study is the selection and evaluation of models used in medico-biological research, such as the torque-angle relationships of the extensors of the elbow; maximum oxygen consumption VO_{2max} ; and the relationship between the results of docking and biological studies in ligand-receptor interactions.

The objectives of this study are: (1) computer modeling and selection of “optimal” models with different optimality criteria for objects related to medical-biological investigations and (2) development of appropriate software programs for routine application of these investigations.

To achieve these objectives we have resolved the following problems:

- Examined and analyzed the current state of modeling and optimization, and their application in the field of medico-biological research;
- Selected and implemented appropriate methods: least squares (LS), robust regression (RR) and minimax fitting; and suitable software for evaluation and modeling of the obtained experimental data;
- Selected and implemented appropriate selection criteria for the “optimal” model from different class models described torque-angle relationships: Akaike information criterion (AIC), Bayesian information criterion (BIC) and root mean square deviation (RMSD); and proposed new criteria: minimizing the maximum residual (MMR) and Hausdorff distance criterion (HDC));
- Developed software programs to calculate the criteria for selecting the “optimal” model and the VO_{2max} ;
- Using correlation analysis, we found dependence between the results of docking and the biological research and thus assessed the used software's for docking (MVD and GOLD);
- Formulated conclusions based on the Interpretation of the results obtained by applying a consistent strategy in the analysis and modeling of experimental data in medico-biological research.