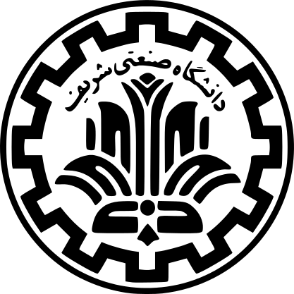
**باسمه تعالی**

****

**دانشگاه صنعتی شریف**

**دانشکده مهندسی برق**

**درس علوم اعصاب محاسباتی**

**گزارش پروژه درس**

**سری دوم**

**یافتن معیاری برای عمق خواب مبتنی بر سیگنال‌های EEG**

**استاد درس:**

**دکتر کربلایی آقاجان**

**امیرحسین افشارراد**

**95101077**

**طاها انتصاری**

**95101117**

**بهار 1397**

# قسمت اول) آشنایی با مقاله پژوهش اصلی

### Terminology:

### Slow-wave=موج آرام(فرکانس پایین)=:SW

### Slow Wave Power=انرژی موج آرام(فرکانس پایین)=:SWP

### Slow Wave microcontinuity =مقداری از انرژی موج آرام که در آینده نزدیک 0.02ثانیه ادامه دارد=:SW%

### Non Rapid Eye Movement=:NREM

### Electroencephalogram=:EEG

### 1- با مطالعه abstract و introduction، هدف کلی این مقاله را توضیح دهید.

هدف این پژوهش یافتن رابطه بین گین مدار فیدبک SW و اثر **سن فرد**، **زمان شب** و **داروی خواب** می‎باشد. همچنین این پژوهش در صدد بررسی این است که آیا گین فیدبک میتواند نمایانگری از اثر جنسیت بر روی SWP باشد یا نه. پژوهش ها نمایانگر این می‎باشند که SW% بهتر میتواند عمق خواب را تخمین بزند.

در این پژوهش تاکید بر روی استفاده از SW% با دلیل عدم تاثیرپذیری این ویژگی سیگنال از ویژگی های آناتومی بدن و جنسیت فرد می‎باشد.همچنین روشSWP هم در حالت های ممکن به صورت "نسبی" پیاده سازی شده است .اما استفاده از روش نسبی SWP به خاطر نبود دلیل دقیق فیزیولوژی پشت بسیاری از این مقادیر، قابل استناد نمی‎باشند.

### 2- آزمایش انجام‎شده برای این پژوهش را توضیح دهید.

برای آزمایش 22 نفر مرد و زن از بازه های سنی متفاوت انتخاب شده اند. داده از افراد در طی 2 شب گرفته شده است. هر فرد در یکی از شب ها از داروی خواب temazepam استفاده کرده است.

در روز بعدی از افراد نیز در مورد کیفیت خوابشان سوال هایی پرسیده شده است که این پرسش ها برای بررسی برداشت خود افراد از خواب با تحلیل های حاصل از پردازش سیگنال ها برای بررسی اثر داروی خواب می‎باشد.که البته جواب این سوال منفی می‎باشد و تا کنون همبستگی بین برداشت خود فرد از خواب و طول کلی خواب NREM مشاهده نشده است. البته تاثیر داروی خواب را میتوان در تفاوت های موجود در SW% دید.

### 3- تعبیر مقاله از عمق خواب را توضیح دهید. همچنین تفاوتش را با تعاریف پیشین، و ایراد آن تعاریف را توضیح دهید.

پژوهش با مدل کردن تولید موج آرام توسط شکل موجود در سوال بعد، عمق خواب را متناظر با ضریب فیدبک شبکه نورونی بیان میکند. در واقع عمق خواب متناظر است با افزایش فیدبک شبکه نورونی که باعث تولید SW میشود.

نتایج پژوهش بیانگر این است که SW% ،با توجه به این که بر پایه مدلی فیزیولوژیک از خواب می‎باشد، عمق خواب را بهتر از SWP نشان میدهد.

از آنجایی که تغییر عمق خواب با تغییرات سنکرون در دامنه فرکانس‎های پایین سیگنال EEG همراه می‎باشد، بسیاری از پژوهش های پیشین از انرژی موج فرکانس کم(SWP) استفاده کرده اند. این در حالی می‎باشد که SWP از آناتومی بدن تاثیر میپذیرد. همچنین جنسیت افراد تاثیر بسزایی در مقدار این ویژگی دارد به طوریکه زن ها دارای مقدار 2برابر مقدار مرد ها میباشند در حالی که اثبات شده است که این نمایانگر عمق خواب افراد نیست(ارجاع اثبات به مقالات ارجاعی 5 و 6). همچنین این عدم قطعیت هنگامی که اثرات "درون فردی" بر روی کل زمان خواب NREM مطالعه میشوند نیز موجود می‎باشد.محققان با ذکر اینکه امکان گریز از این مشکلات با بررسی نسبی مقادیر SWP وجود دارد این نکته را بیان میکنند که این مقادیر نسبی پایه فیزیولوژی ندارند.

جدا از توضیح بالا مبنی بر عدم وجود علت برای برخی از این مقادیر نسبی، به دلیل اینکه این مقادیر نسبی در حوزه فرکانس حساب میشوند، عدم دقت زمانی کافی باعث کم شدن دقت و در نتیجه تخمین بد از جواب حقیقی میشود. البته این دقت زمانی کم برای روش گین فیدبک تاثیری ندارد چرا که این مدل تماما در حوزه زمان پیاده سازی میشود.

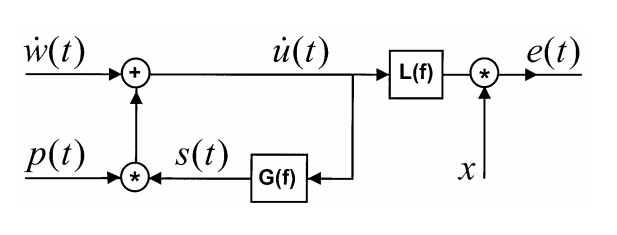
### 4- قسمت دوم مقاله (THE FEEDBACK MODEL)را به طور کلی مطالعه کنید. (ممکن است بعضی از مفاهیم به کل برایتان جدید باشد، برای همین نیازی نیست که دقیق شوید). کلیت مدل پیشنهادی را توضیح دهید. دلیل نویسندگان مقاله برای انتخاب این مدل چه بوده است؟ به طور شهودی ارتباط این مدل با عمق خواب و سیگنال EEGرا توضیح دهید

با توجه به ویژگی های فیزیولوژیکی، مدل ما بایستی 3 ویژگی را داشته باشد:

1. اینکه مقدار گین فیدبک با عمق خواب متناسب می‎باشد.

2.اثرات تصادفی و غیر پیش بینی شده بیرونی که مدل را تحت تاثیر قرار میدهند را در نظر بگیرد

3. گین فیدبک بزرگتر متناظر است با اینکه SWP بزرگتر باشد.



مدلی که پیاده سازی شده است به صورت بالا می‎باشد.در مدل بالا هر اثر نویز با ورودی  که به صورت نویز سفید گوسی فرض شده است مدل میشود.

فیلتر G(f) که موج آرام را از خروجی استخراج میکند و پس از آن با ضرب شدن این مقدار در مقدار گین فیدبک p(t)، فیدبک نهایی مدل آماده می‎باشد. p(t) نمایانگر عمق خواب می‎باشد که مشخص میکند که چه کسری از سیگنال کنونی موجود در مغز، در آینده نزدیک ادامه می یابد.

سیگنال موجود در مغز پس از عبور از یک فیلتر پایین گذر L(f) و یک ضریب نامرتبط به خواب، x،(ویژگی های آناتومی سر) در الکترود های EEG ضبط میشود.

در مدل بالا تابع G(f) یک فیلتر میان گذر خطی می‎باشد که مطابق آنچه که در صورت مقاله ذکر شده تعریف میشود. برای این پژوهش مورد نظر فرکانس های قطع پایین و بالا به ترتیب 0.5 و 2 هرتز میباشند در مورد این فیلتر ذکر شده است که تنها با شیب 6db/oct افت میکند. همانگونه که مقاله توضیح میدهد، علت عدم افت شدید اندازه این فیلتر، ماهیت خود داده های موج آرام می‎باشد، چرا که در بر داشتن اطلاعاتی از محدوده‎های فرکانسی مجاور نیز می‌تواند حاوی اطلاعات مفیدی باشد.

این مدل شبکه نورون را میتوان در ارجاعات 6 و 10 مقاله یافت:

6. M. S. Mourtazaev, B. Kemp, A. H. Zwinderman, and H. A. C. Kamphuisen, “Age and gender affect different characteristics of slow waves in the sleep EEG,”

10. Accurate measurement of flash-evoked alpha attenuation [B.Kemp](https://www.sciencedirect.com/science/article/pii/0013469483900792?via%3Dihub#!)

به نظر میرسد از اولین استفاده های این مدل فیدبک در ارجاع فوق (شماره 10) صورت گرفته باشد.

# قسمت دوم) آشنایی با دیتاست

## 2-2)

تابع FeatureExtraction را نوشتیم که فایل آن در کنار فایل اصلی ضمیمه شده است. ورودی‎ و خروجی‎های این تابع به شرح زیر است:

**خروجی‎ها:**

* t : زمان آغاز هر بازه 10 ثانیه‎ای
* state : وضعیت خواب در طول هر کدام از بازه‎های 10ثانیه‎ای
* X : ماتریسی با 10 سطر و N ستون که N تعداد بازه‎های 10ثانیه‎ای مربوط به سیگنال اصلی می‎باشد. سطرهای اول تا چهارم انرژی باندهای دلتا، تتا، آلفا، و بتای مربوط به الکترود Fpz و سطرهای پنجم تا هشتم مشابه چهار سطر اول، این بار مربوط به الکترود Oz هستند که هر درایه توان متوسط مربوط را در بازه 10ثانیه‎ای متناظر نشان می‎دهد. دو سطر آخر نیز توان متوسط مربوط به سیگنال‎های EOG و EMG می‎باشند.

**ورودی:**

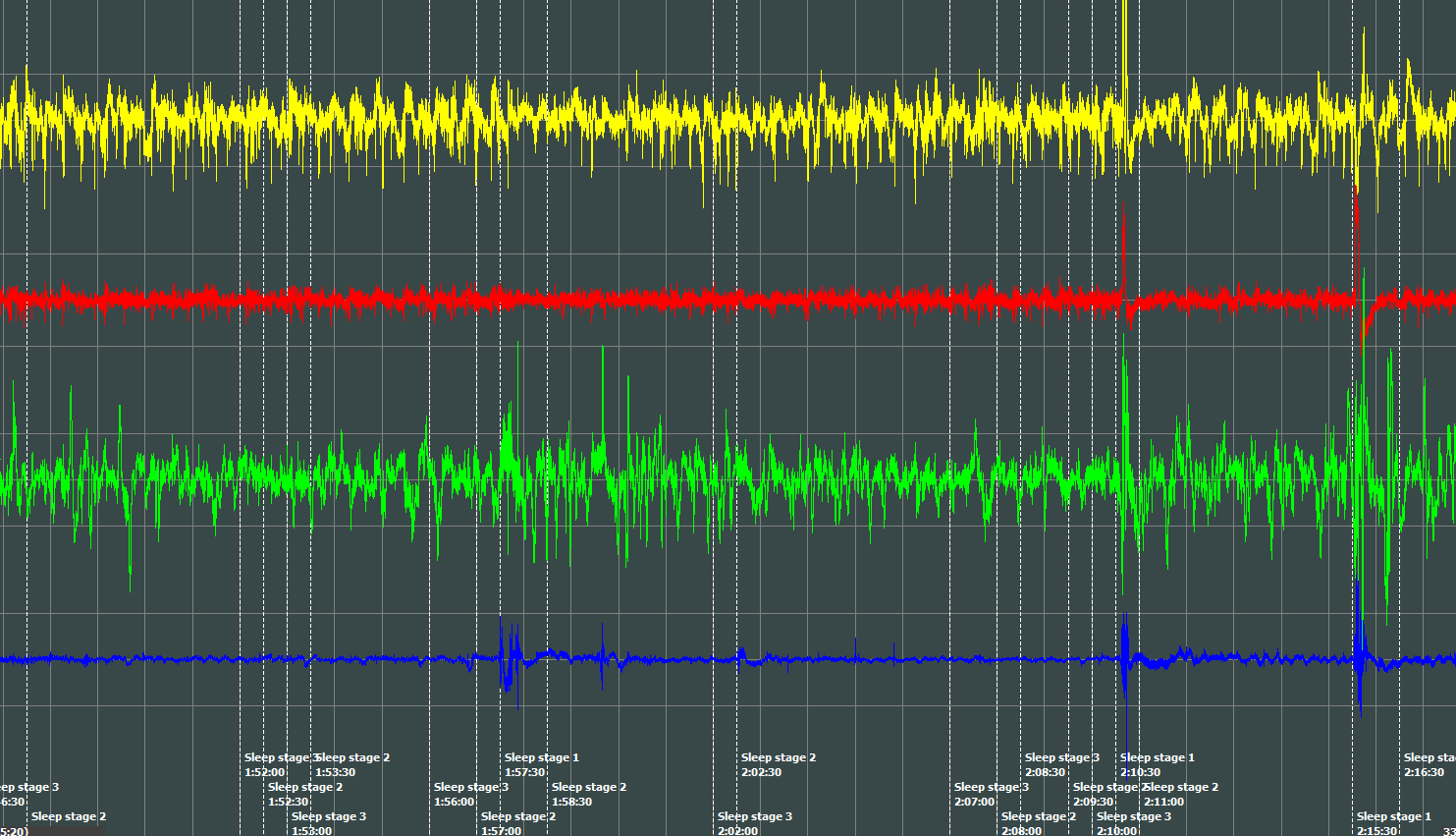
* edf\_path: آدرس فایل edf مربوطه می‎باشد.
* hyp\_path: آدرس فایل هایپنوگرام مربوطه می‎باشد.

**نکته1:** برای محاسبه توان متوسط سیگنال در یک بازه، از رابطه‎ی زیر استفاده شده است:

**نکته2:** برای انتخاب طول پنجره‎های زمانی، گزینه‎های متعددی وجود دارد. برای انجام این انتخاب، باید سعی کرد به گونه‎ای عمل کرد که هم نتایج خوبی بگیریم و هم تا حد امکان، دشواری بی‎مورد به کار خود اضافه نکنیم. با بررسی زمان‎های تغییر وضعیت خواب (خروجی تابع AnnotExtract)، می‎بینیم که اولاً همه‎ی زمان‎ها مضارب 10 هستند، و ثانیاً عدد بزرگ‎تری وجود ندارد که همه‎ی این زمان‎ها بر آن بخش‎پذیر باشند، لذا با انتخاب یک عدد دیگر، برخی از زمان‎های تغییر وضعیت خواب جزء مضارب صحیح آن عدد نخواهند بود و این عملکرد کار ما را دچار مشکل می‎کند. از طرفی، انتخاب شمارنده‎های 10، مانند 2 یا 5 نیز اگرچه امکان‎پذیر است، اما ممکن است کیفیت کار ما را دچار مشکل کند، چرا که اگر پنجره‎های زمانی بیش از حد کوچک شوند، محتوای انرژی آن‎ها ممکن است فاقد اعتبار لازم باشد و به دلیل کوچک بودن بازه زمانی، مقادیر انرژی باندهای مختلف و مقایسه‎ی آن‎ها با یکدیگر، نتیجه‎ی دقیقی به ما ندهد. بنابراین عدد 10 را به عنوان طول پنجره زمانی انتخاب می‎کنیم. (در سوال پنجم، این عدد را تغییر می‎دهیم تا ببینیم آیا کوچک کردن این بازه، تأثیری -چه مثبت، چه منفی- روی عملکرد ما خواهد داشت یا نه.)

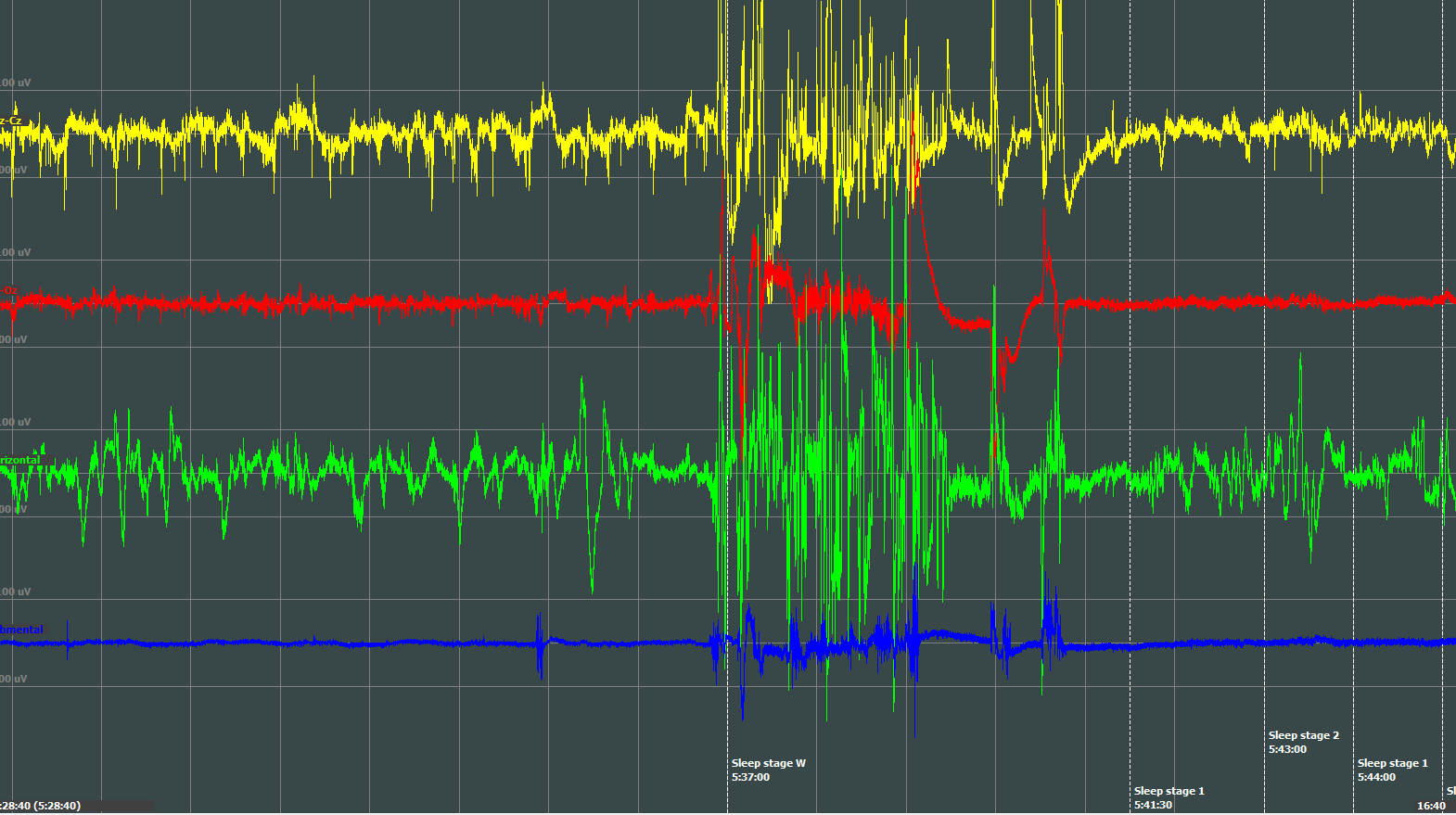
## 2-3)

به طور کلی، محیط برنامه به شکل زیر است:



که البته مستقیماً از روی شکل‎های زمانی سیگنال‎ها، نمی‎توان اطلاعات چندانی به دست آورد. در این جا، یک نکته‎ی استنباط شده از سیگنال‎های زمانی را ذکر می‎کنیم، و سپس به ارائه اطلاعاتی از مراجع دیگر می‎پردازیم.

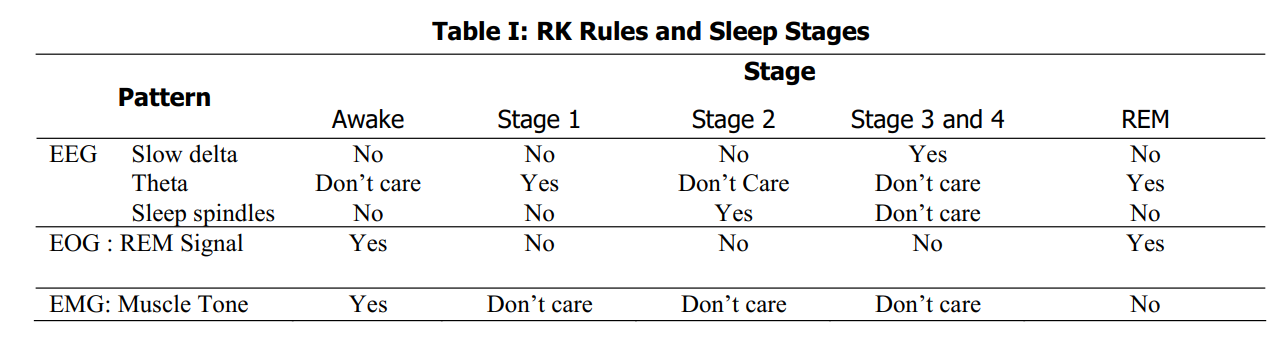
به استناد به سیگنال‎های مشاهده شده در نرم افزار، هرگاه تغییر حالت از یکی از سطوح خواب به بیداری صورت بگیرد، در تمامی سیگنال‎ها، نوسانات پردامنه، شدید و نامنظمی مشاهده می‎شود که می‎تواند ملاک بسیار دقیقی برای تشخیص زمان ورود به حالت بیداری از طریق سیگنال‎های زمانی باشد. نمونه‎ای از این سیگنال‎ها را در شکل زیر مشاهده می‎کنید:



در مورد سیگنال‎های دریافتی، معمولاً از سیگنال EOG برای شناسایی حرکت پلک‎ها (حالت‎های REM - و SEM -) استفاده می‎شود.

از سیگنال EMG معمولاً برای شناسایی حرکات عضلانی، و در نتیجه برای شناسایی حالت بیداری استفاده می‎شود.

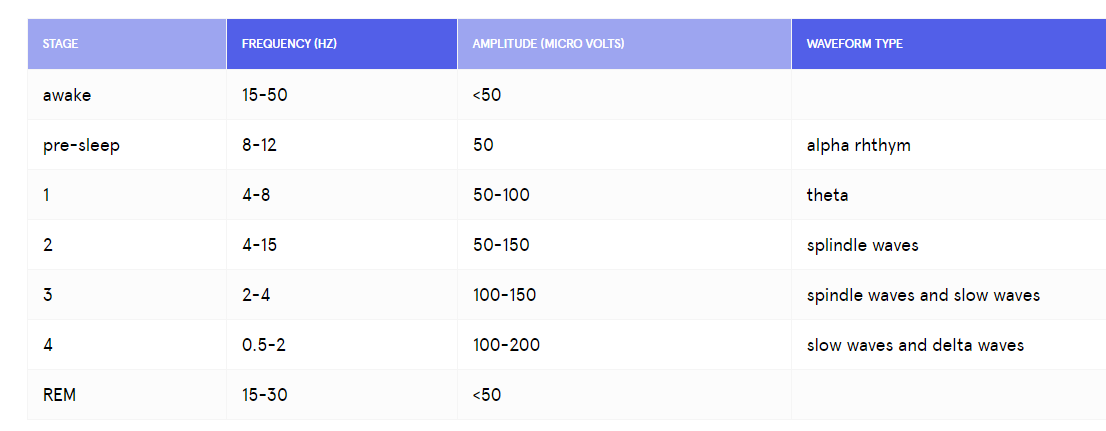
از سیگنال EEG برای استخراج اطلاعات 4 لایه‎ی خواب استفاده می‎شود، به طوری که باند دلتا عموماً برای لایه‎های 3 و 4 (در ادامه این پروژه، این مسأله تأیید می‎شود)، و به عنوان مثال از sleep spindle های موجود در این سیگنال‎ها برای شناسایی لایه 2 و نیز از باند تتا برای لایه اول استفاده می‎شود. البته روش یکتایی برای ارتباط برقرار کردن بین سیگنال‎ها و لایه‎های خواب وجود ندارد و می‎توان ویژگی‎های دیگری نیز ارائه کرد. آن چه گفته شد در جدول زیر خلاصه شده است:



مرجع:

EOG and EMG: Two Important Switches inAutomatic Sleep Stage ClassificationE Estrada, H Nazeran, J Barragan , J R Burk\*, E A Lucas\*, and K Behbehani\*\*  
Department of Electrical and Computer Engineering, University of Texas at El Paso, El Paso TX, USA  
\* Sleep Consultants Inc, Ft Worth Texas, USA  
\*\*  
Department of Bioengineering, University of Texas at Arlington, Arlington, Texas, USA  
E-mail: nazeran@ece.utep.edu

یک تقسیم‎بندی دیگر برای اطلاعات EEG و مراحل خواب در جدول زیر آمده است:



مرجع : <https://www.tuck.com/stages/>

# قسمت سوم) یافتن معیاری برای توصیف خواب مبتنی بر PCA

## 3-1)

PCA روشی برای کاهش ابعاد فضا با از دست دادن کمترین اطلاعات ممکن است. در این روش، ابتدا ماتریس کوواریانس متغیرهای تصادفی موجود محاسبه می‎گردد. (تعداد ابعاد فضا، N، همان تعداد متغیرهای تصادفی است، که در هر مسأله با توجه به نمونه‎گیری انجام شده، تعدادی از تحقق‎های این متغیر تصادفی - تعدادی نقطه در این فضای N بُعدی - در اختیار است)

سپس با محاسبه مقادیر و بردارهای ویژه‎ی ماتریس کوواریانس، می‎توان راستاهایی در فضا را یافت که اطلاعات در آن راستاها بیشترین پخش‎شدگی را داشته باشند. واریانس در هر راستا، همان مقادیر ویژه هستند و بردارویژه‎ی متناظر با هر مقدار ویژه، نشان‎دهنده‎ی راستای متناظر در فضای مذکور است. بنابراین کافی است مقادیر ویژه را به ترتیب نزولی مرتب کنیم و بردارویژه‌های متناظر را به عنوان جهت‌های ماکزیمم تغییرات معرفی کنیم. نهایتاً می‎توان حد آستانه‎ای (مثلاً 90درصد) را در نظر گرفت و تنها راستاهایی از فضا را نگه داشت که مجموع مقادیر ویژه (واریانس‎ها) برای آن‎ها، بیش از حد آستانه‎ی معین‎شده باشد. عموماً مشاهده می‎شود که در نظر گرفتن 3 راستا، کافی است. با این کار، ابعاد فضا را کاهش داده‎ایم و کمتر از 10 درصد اطلاعات از دست رفته است.

نهایتاً باید نقاط موجود در فضا را روی این 3 (یا هر تعداد انتخابی دیگری) تصویر کنیم و اطلاعات را در فضای جدید (با بُعد کاهش‎یافته) منعکس کنیم.

خروجی‎های تابع pca عبارتند از:

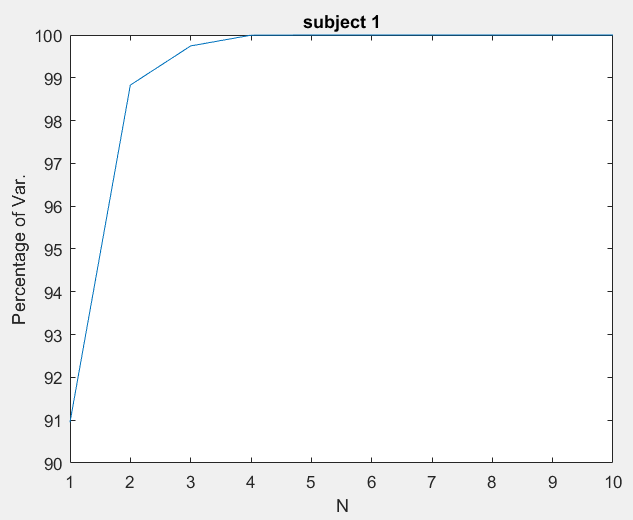
**coeff:** ماتریسی است که تمامی بردارهای ویژه‎ی راستاهای استخراج‎شده توسط PCA را به صورت ستونی و به ترتیب اهمیت (به ترتیب نزولی مقدار ویژه) در بر دارد.

**score**:ماتریسی است که ابعاد آن، با ابعاد ماتریس ورودی تابع یکسان است و نقاط ورودی را در همان فضای اولیه (بدون کاهش بُعد) اما با یک پایه‎ی متعامد جدید (که همان بردارهای ویژه هستند - می‎دانیم که به دلیل متقارن بودن ماتریس کوواریانس، این بردارها متعامد هستند و با در نظر گرفتن تمامی آن‎ها، پایه‎ای متعامد برای فضای اولیه خواهیم داشت -) توصیف می‎کند. بنابراین همان نقاط را در این ماتریس داریم، اما مختصات آن‎ها عوض شده است، چرا که پایه‎ی فضا را تغییر داده‎ایم.

**latent:** برداری به طول N است که مقدار ویژه‎ی متناظر با بردارهای ویژه‎ی موجود در ماتریس coeff را در بر دارد. مقادیر ویژه و بردارهای ویژه‎ی متناظر به ترتیب نزولی (به ترتیب کاهش اهمیت) مرتب شده‎اند.

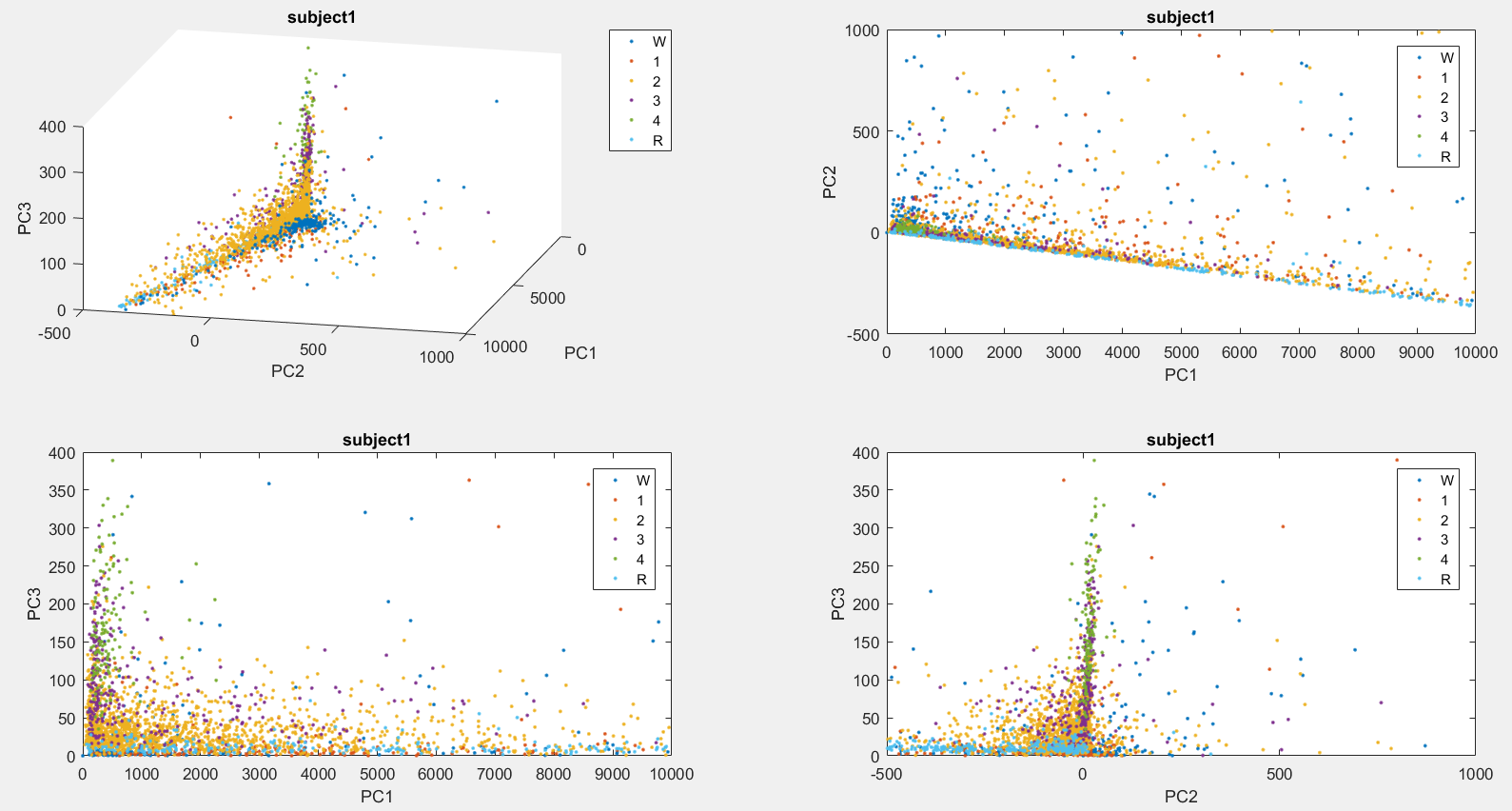
## 3-2)

نمودار تجمعی واریانس (انرژی) برای سابجکت اول به صورت زیر است:



مطابق نمودار فوق، حدود 91 درصد انرژی در راستای اول، حدود 98.8 درصد انرژی در دو راستای اول، حدود 99.8 انرژی در سه راستای اول، و برای مابقی راستا‎ها، تقریباً تمامی انرژی در n راستای اول ذخیره شده است. این مشاهده با توضیحات بخش قبل نیز سازگار است و بیان می‎کند که انتخاب 3 راستای اول می‎تواند برای کاهش بُعد فضا و از دست ندادن اطلاعات، انتخاب بسیار مناسبی باشد.

## 3-3 و 3-4)



نمودارهای شکل فوق، شامل نمودار سه بعدی خواسته شده، و نیز تصویر نقاط در سه صفحه‎ی مختصات این نمودار هستند.

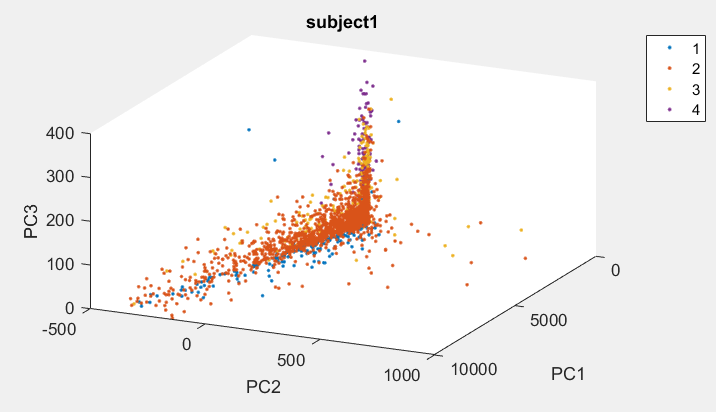
(تذکر: مقیاس‎های محورهای مختصات یکسان نیستند)

با توجه به نمودار فوق، مشهود است که نمی‎توان ادعا کرد هیچ کدام از این سه راستا یا ترکیبات آن‎ها قادر به جداسازی همه‎ی 6 حالت موجود هستند. با این وجود می‎توان مشاهده نمود که راستای PC1 تا حد قابل قبولی، همه‎ی دیتای مربوط به خواب REM را در بر دارد. (البته با توجه به نمودار رسم شده در صفحه‎ی PC1-PC2، به شکل دقیق‎تر می‎توان گفت که راستایی که اطلاعات REM را در بر دارد، خطی در این صفحه است که از مبدأ گذشته و زاویه‎ای در حدود -30 درجه دارد)

همچنین می‎توان ادعا نمود که وضعیتی نسبتا مشابه برای حالت بیداری نیز موجود است ونقاط مربوط به این وضعیت، اکثراً در صفحه‎ی PC1-PC2 پخش شده‎اند و راستای PC3، حجم کمی از این نقاط را در بر می‎گیرد.

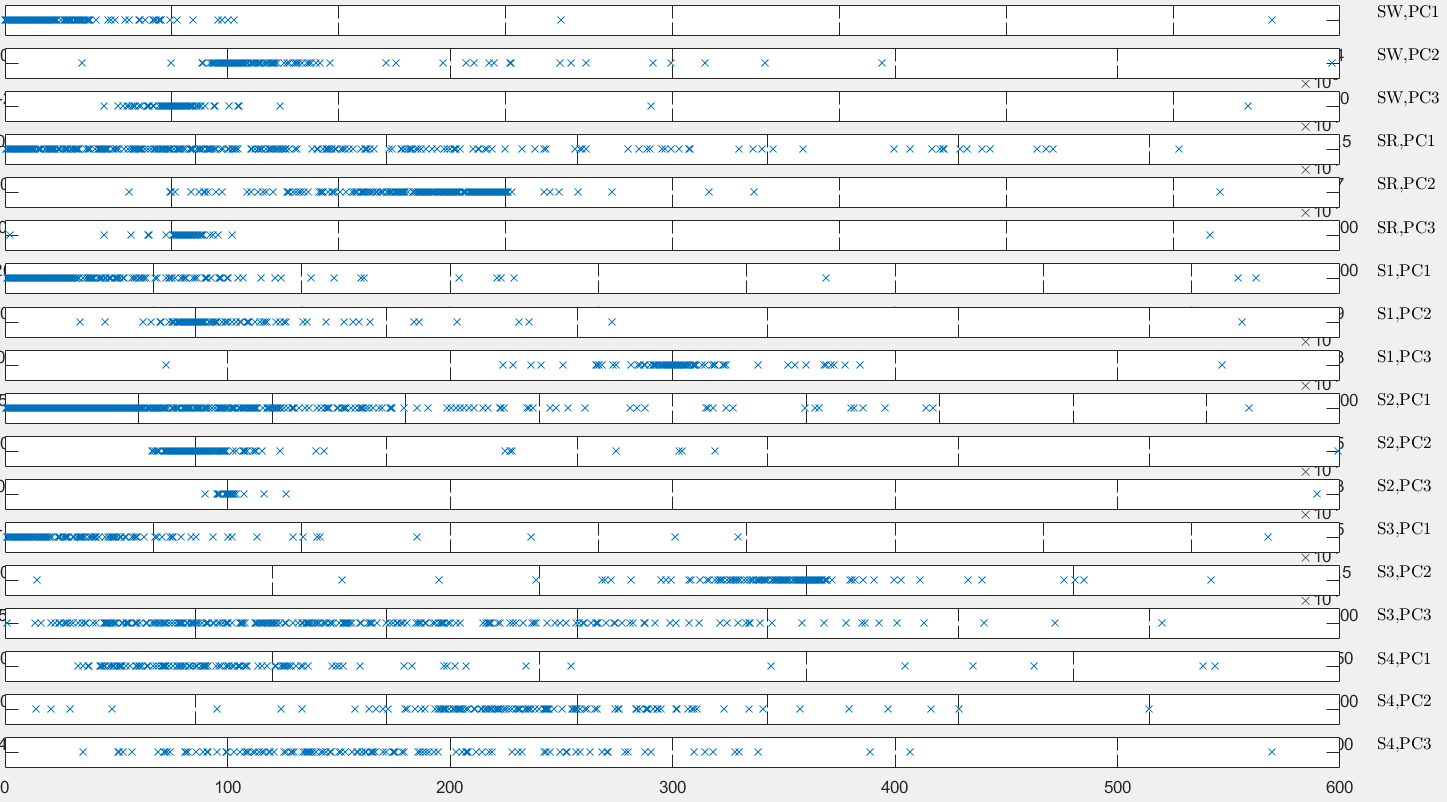
همچنین با دقت به راستای PC3 در نمودارهای مختلف، درمی‎یابیم که محتویات این راستا، عمدتاً مربوط به خواب NREM (و به طور خاص، بیشتر شامل عمق 3 و 4) است و این راستا، عمق‎های مختلف NREM را می‎تواند تا حد قابل قبولی از هم جدا کند.

برای بررسی دقیق‎تر وضعیت‎های مختلف خواب NREM، می‎توانیم نمودار مربوط به این 4 وضعیت را رسم کنیم.



آن‎چه مشخص است، راستای PC2 اطلاعات چندانی در بر ندارد، و دو راستای PC1 و PC3 در تمایز دادن داده‎های مربوط به لایه‎های مختلف NREM موفق‎تر هستند.

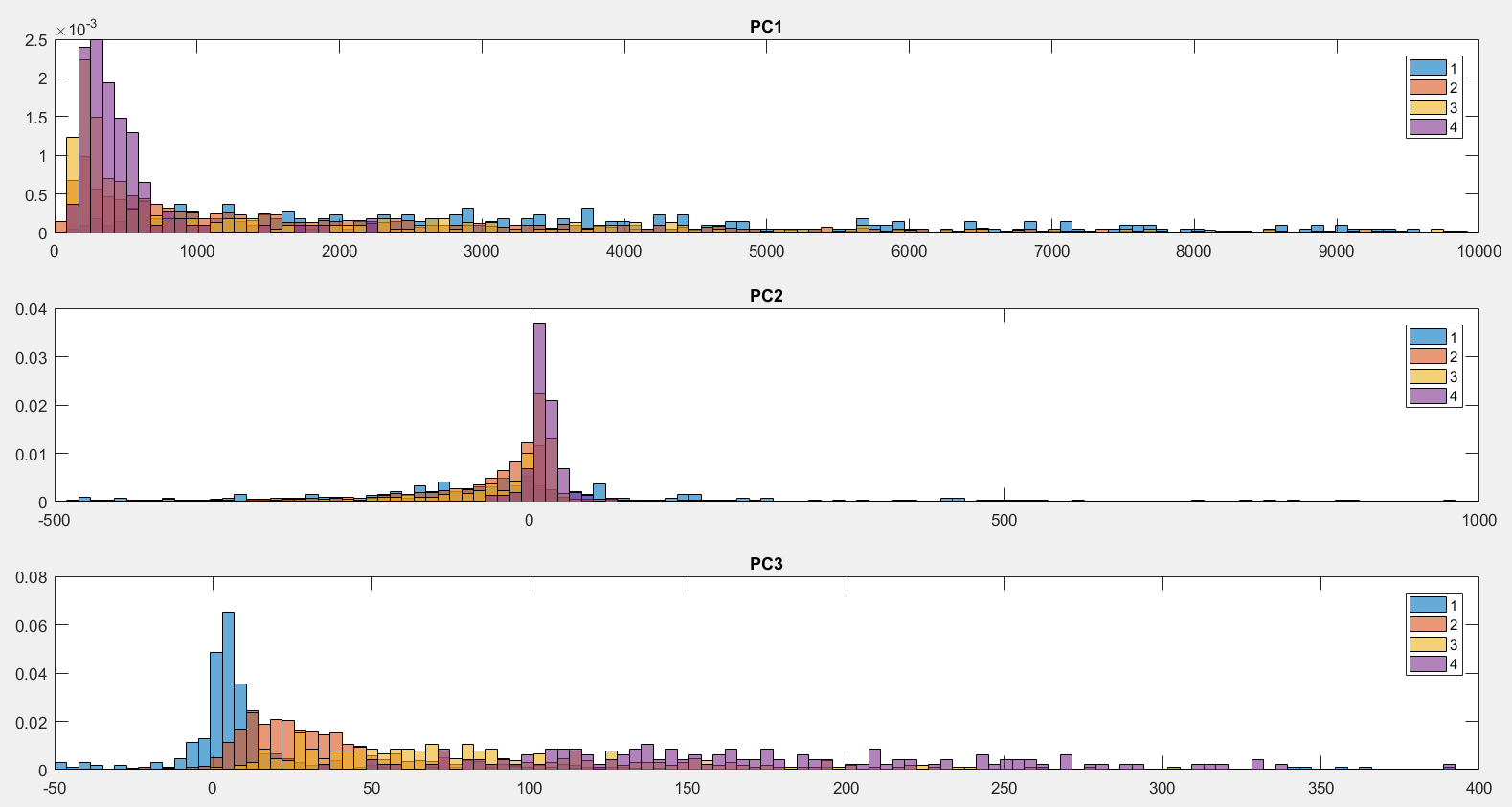
همچنین می‎توان در نمودارهایی یک‎بعدی، توزیع نقاط مربوط به هر یک از 6 لایه را در راستای PC1، PC2، و PC3 (مجموعاً 18 نمودار) مشاهده کرد. این نمودارها به شکل زیر هستند:



با توجه به نمودارهای فوق، می‎توان مشخص کرد که هر کدام از 6 لایه‎ی خواب، در کدام یک از راستاها کشیدگی بیشتری دارند. به عنوان مثال، وضعیت REM، لایه دوم، لایه سوم، و لایه چهارم به ترتیب در راستاهای PC1، PC1، PC3 و PC3 کشیدگی قابل توجهی دارند.

از طرف مقابل به عنوان مثال می‎توان مشاهده کرد که راستای PC3 برای لایه دوم، و نیز وضعیت REM تقریباً هیچ اطلاعاتی در بر ندارد.

اطلاعات موجود در نمودار قبل را می‎توان در قالب هیستوگرام نیز بیان کرد، که می‎تواند شهود بهتری از محل‎های تراکم نقاط در وضعیت‎های مختلف خواب در اختیار ما قرار دهد. هیستوگرام مذکور برای حالت‎های NREM به شکل زیر می‎باشد: (سه هیستوگرام برای سه راستای PC1، PC2، و PC3)



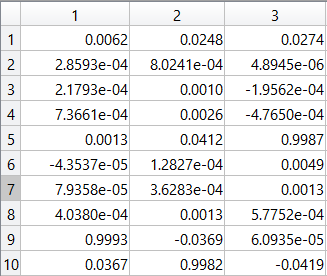
می‎بینیم که لایه‎های 1، 2، 3، و 4 به ترتیب بیشترین کشیدگی را در راستاهای PC1، PC1، PC3 و PC3 دارند. در این جا کشیدگی اطلاعات به یکنواخت بودن هیستوگرام و نداشتن قله تعبیر می‎شود. نتایج مذکور با آن‎چه از تحلیل نمودارهای قبلی به دست آمد، منطبق است.

یک نکته‎ی مهم نیز آن است دو دسته از نقاط می‎توانند توسط یک راستا از هم جدا شوند، حتی اگر در آن راستا کشیدگی بسیار زیادی نداشته باشند. به عنوان مثال، با دقت به هیستوگرام سوم می‎بینیم که تجمع 4 حالت خواب در بخش‎های مختلفی اتفاق افتاد است، و این بدان معناست که راستای PC3 در جداسازی 4 حالت NREM از یکدگیر نسبتاً موفق عمل کرده است.

در پایان، تکرار یکی از نکات مذکور در توضیحات این بخش خالی از لطف نیست، و آن نکته این است که ما در تحلیل‎های مطرح‎شده در این بخش، خود را محدود به سه راستای PC1، PC2، و PC3 کردیم، و طبعاً این محدودیت باعث می‎شود که نتوانیم داده‎های 6 وضعیت مختلف را به خوبی جدا کنیم. در حالت کلی‎تر، ممکن است اطلاعات مربوط به یک وضعیت، در راستایی از این فضا که به صورت یک ترکیب خطی از سه راستای اصلی قابل بیان است، بیشترین کشیدگی و تمایز را داشته باشد، یا حتی ممکن است در ناحیه‎ای از این فضا، تجمعی شاخص اما غیر خطی داشته باشد. با این حال، با توجه به این که در قسمت بعد، می‎خواهیم با بررسی منشأ سه راستای اصلی، ارتباطی بین سیگنال‎های مختلف EEG و لایه‎های خواب برقرار کنیم، در این بخش به همین تحلیل اکتفا می‎کنیم.

## 3-5)

برای پاسخ به این سؤال، به ماتریس coeff مراجعه می‎کنیم (تنها ستون‎های اول تا سوم این ماتریس در این بخش برای ما اهمیت دارد):



مشاهده می‎شود که در هر سه راستای اول، تنها یکی از ابعاد فضای حالت اصلی نقش تأثیرگذار داشته و سایر ابعاد، تقریباً بی‎اثرند. به تعبیر دیگر، سه راستا از راستاهای فضای حالت اولیه هستند که واریانس بسیار زیادی دارند و می‎توان آن‎ها را به عنوان راستاهای مؤثر انتخاب کرد. با توجه به ضرایب، می‎بینیم که اثر راستاهای دیگر ناچیز است و می‎توان در عمل از آن‎ها صرف نظر کرد.

* راستای اول (PC1) : توان متوسط سیگنال EOG (سایر پارامترها، اثری ناچیز دارند)
* راستای دوم (PC2) : توان متوسط سیگنال EMG (سایر پارامترها، اثری ناچیز دارند)
* راستای سوم (PC3) : توان متوسط باند فرکانسی دلتای الکترود Oz (سایر پارامترها، اثری ناچیز دارند)

با توجه به این که ضریب متناظر برای سه پارامتر معرفی شده در سه راستا، نزدیک به 1 است، در هر سه مورد، مؤلفه‎ی متناظر PCA با افزایش توان متوسط مربوطه افزایش می‎یابد و ارتباطی مستقیم و تقریباً از جنس رابطه‎ی همانی دارند. (اگر این ضریب نزدیک به 1- می‎بود، ارتباط آن‎ها بالا، اما به شکل معکوس بود)

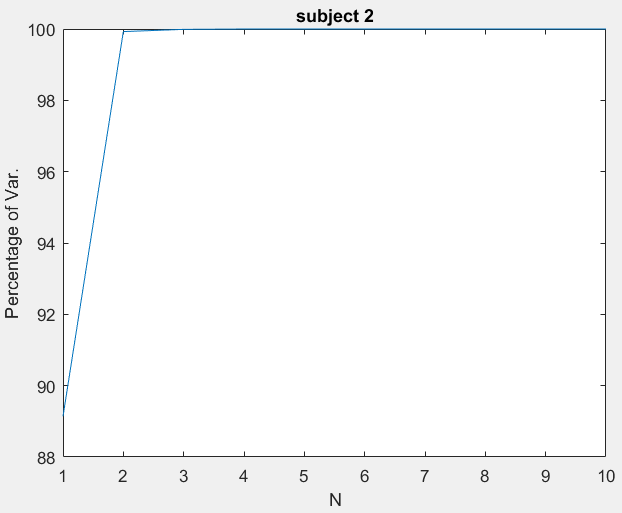
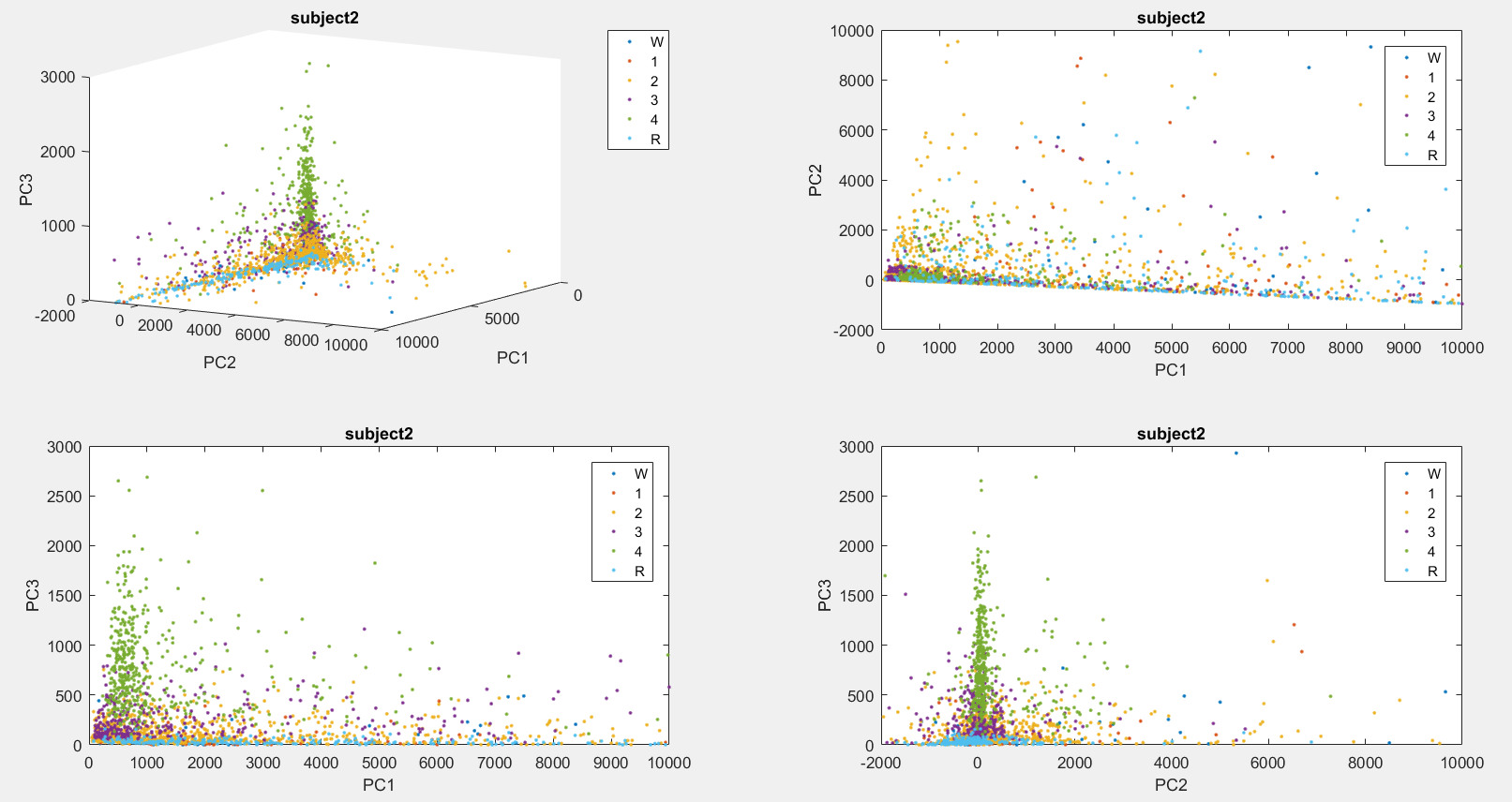
حال، می‎توان با توجه به تحلیل‎های انجام شده در این بخش و بخش قبل، نتیجه‎گیری‎هایی انجام داد:

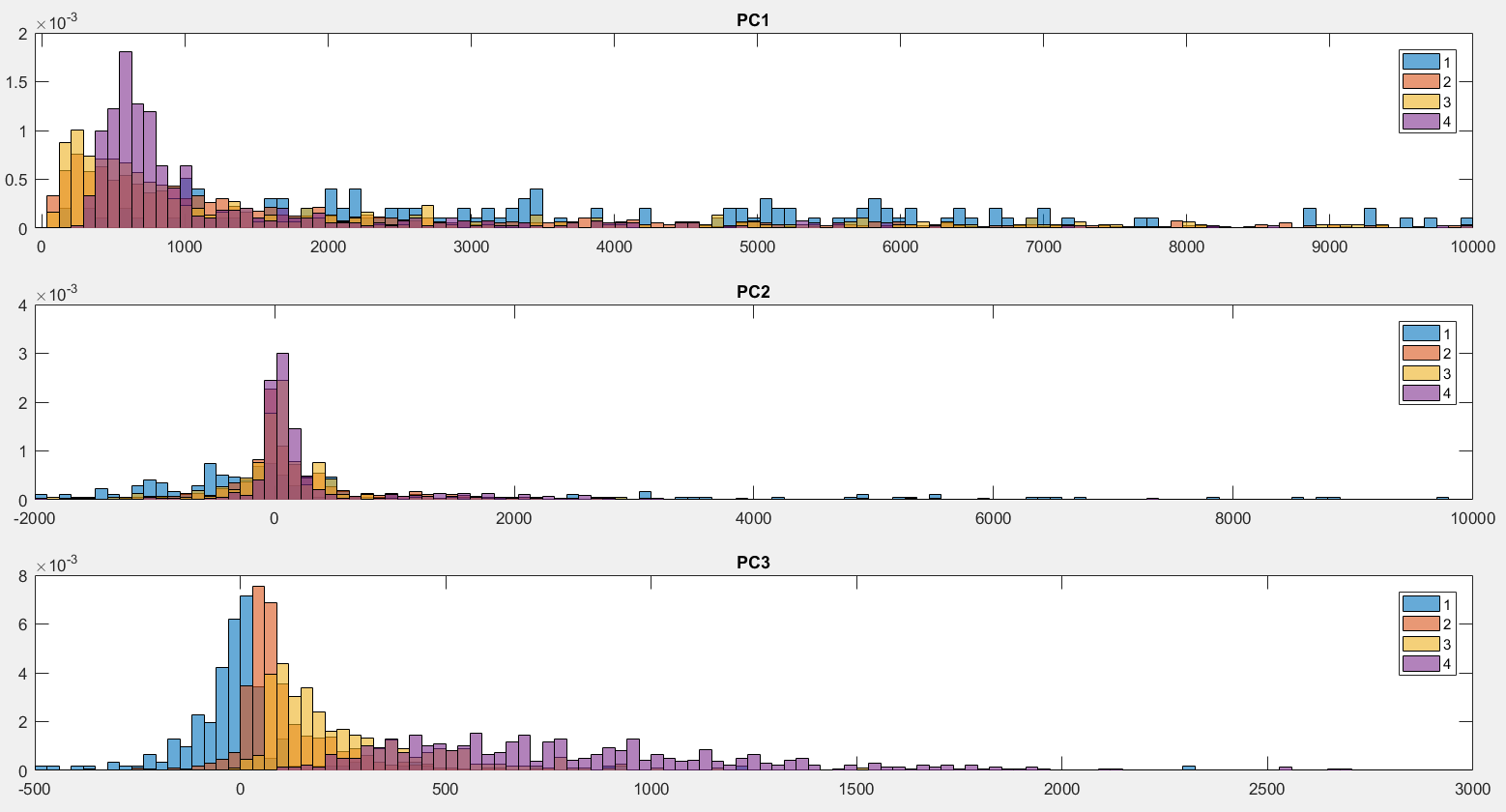
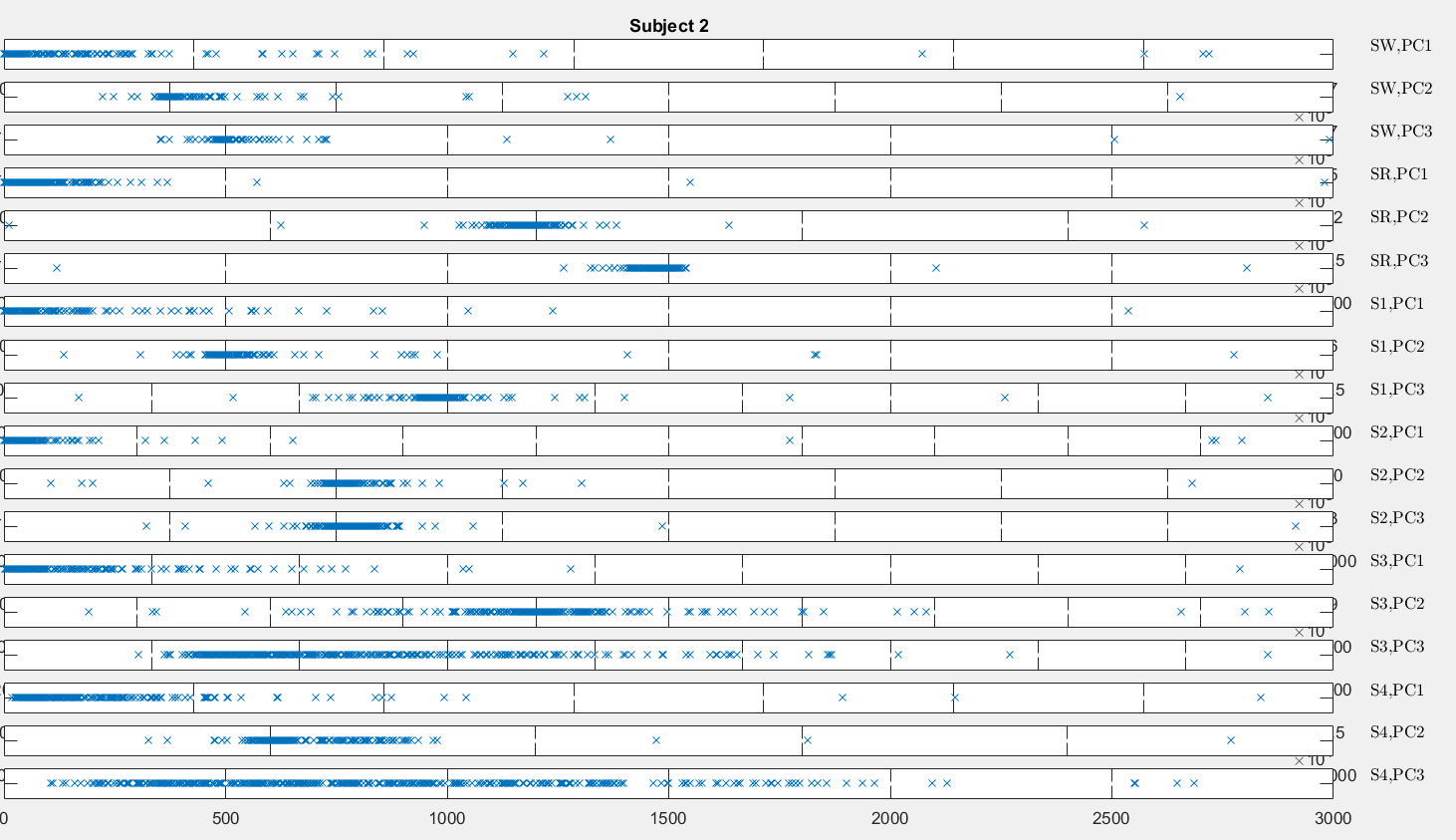
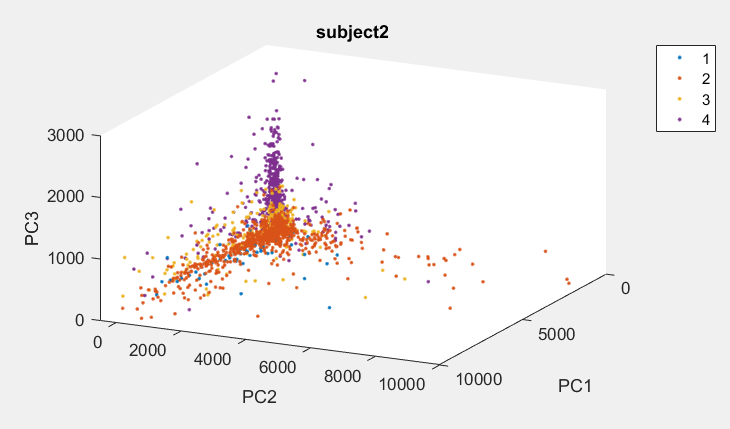
* اعماق مختلف خواب NREM توسط متوسط باند فرکانسی دلتا (مربوط به الکترود Oz) می‎توانند تا حد قابل قبولی از یکدیگر جدا شوند.
* عمق اول و دوم خواب بیشتر در راستای توان متوسط سیگنال EOG دارای کشیدگی هستند و عمق سوم و چهارم بیشتر در راستای باند دلتای الکترود Oz، با این حال با توجه به هیستوگرام‎های رسم شده، باند فرکانسی دلتای الکترود Oz برای جداسازی مجموعه کلی آن‌ها انتخاب مناسب‎تری است.
* توان متوسط سیگنال EMG، اطلاعات چندانی به خصوص در مورد خواب NREM در اختیار ما قرار نمی‎دهد.
* خواب REM در راستایی که ترکیبی خطی از توان متوسط سیگنال EOG و EMG است، توزیع شده و کشیدگی بسیار زیادی دارد. در این راستا، سهم متوسط سیگنال EOG بیشتر است، اما نمی‎توان از اثر سیگنال EMG چشم‎پوشی کرد. بنابراین با در نظر گرفتن این راستا، می‎توان دیتای مربوط به خواب REM را تا حدی از سایرین جدا کرد. (البته با مشاهدات این سابجکت، این کار کیفیت چندانی نخواهد داشت، چرا که اگرچه تمامی دیتای مربوطه به REM در این راستا متمرکز شده است، اما اطلاعات پراکنده‎ای مربوط به سایر وضعیت‎ها نیز در آن موجود است.)

## 3-6)

ابتدا نمودارهای بخش‎های قبل را برای هر کدام از سابجکت‎ها رسم می‎کنیم، سپس با بررسی نمودارها و نیز سه ستون اول ماتریس coeff، نتایج را برای آن‎ها بررسی می‎کنیم. در این بخش، به تفصیل دو بخش قبل برای هر سابجکت صحبت نخواهیم کرد و صرفاً به نتیجه‎گیری می‎پردازیم:

**سابجکت دوم)**

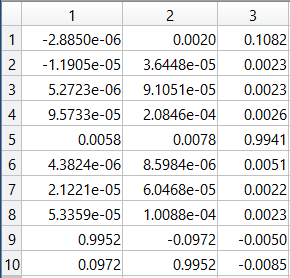
 



آن چه برای این سابجکت مشاهده می‎شود به طور خلاصه آن است که راستای PC3 جداکنندگی قابل توجهی برای اعماق مختلف خواب NREM دارد. این ویژگی در سابجکت قبلی نیز مشاهده شد.

همچنین خواب REM نیز کشیدگی نسبی در راستای PC1 دارد.

همچنین سه بردار ویژه اصلی برای این سابجکت به صورت زیر است:

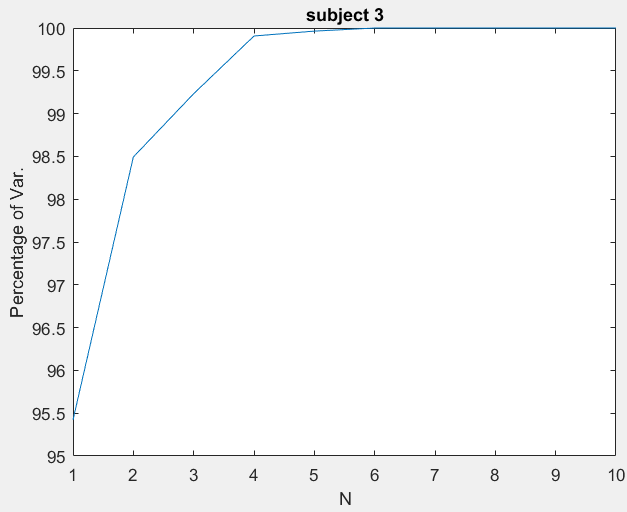


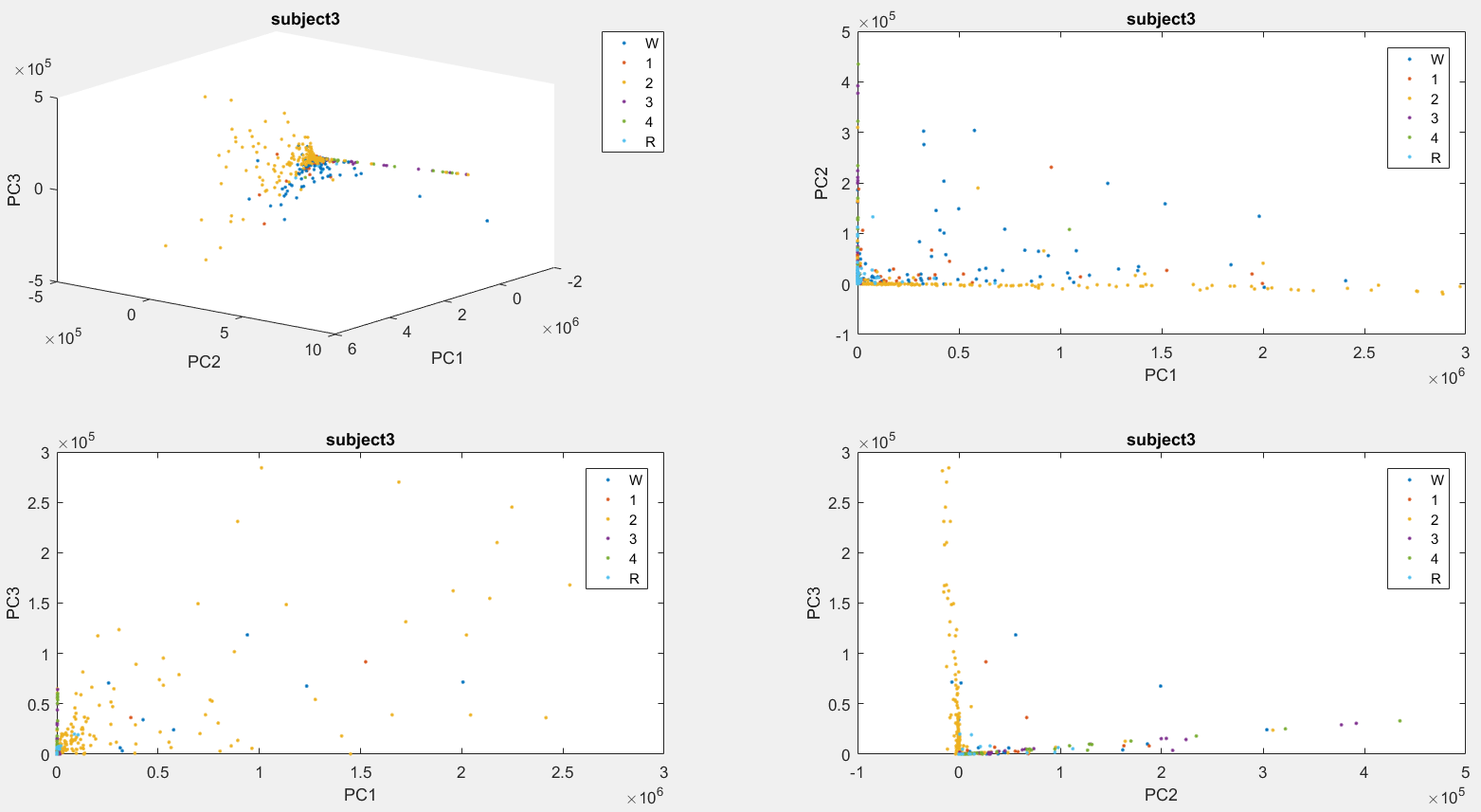
و نتیجه به شرح زیر خواهد بود:

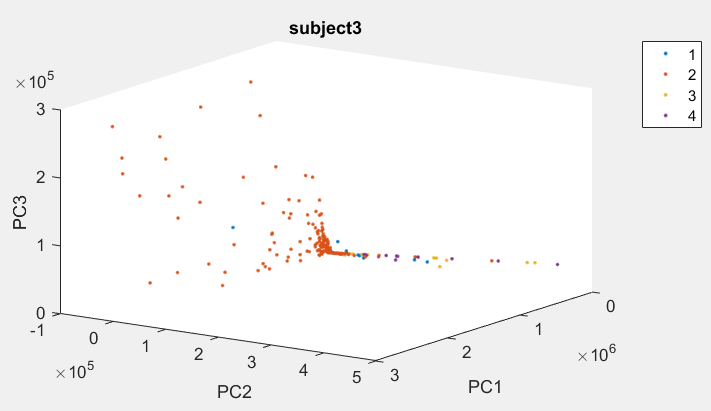
* راستای اول (PC1) : توان متوسط سیگنال EOG (سایر پارامترها، اثری ناچیز دارند)
* راستای دوم (PC2) : توان متوسط سیگنال EMG (سایر پارامترها، اثری ناچیز دارند)
* راستای سوم (PC3) : توان متوسط باند فرکانسی دلتای الکترود Oz (سایر پارامترها، اثری ناچیز دارند)

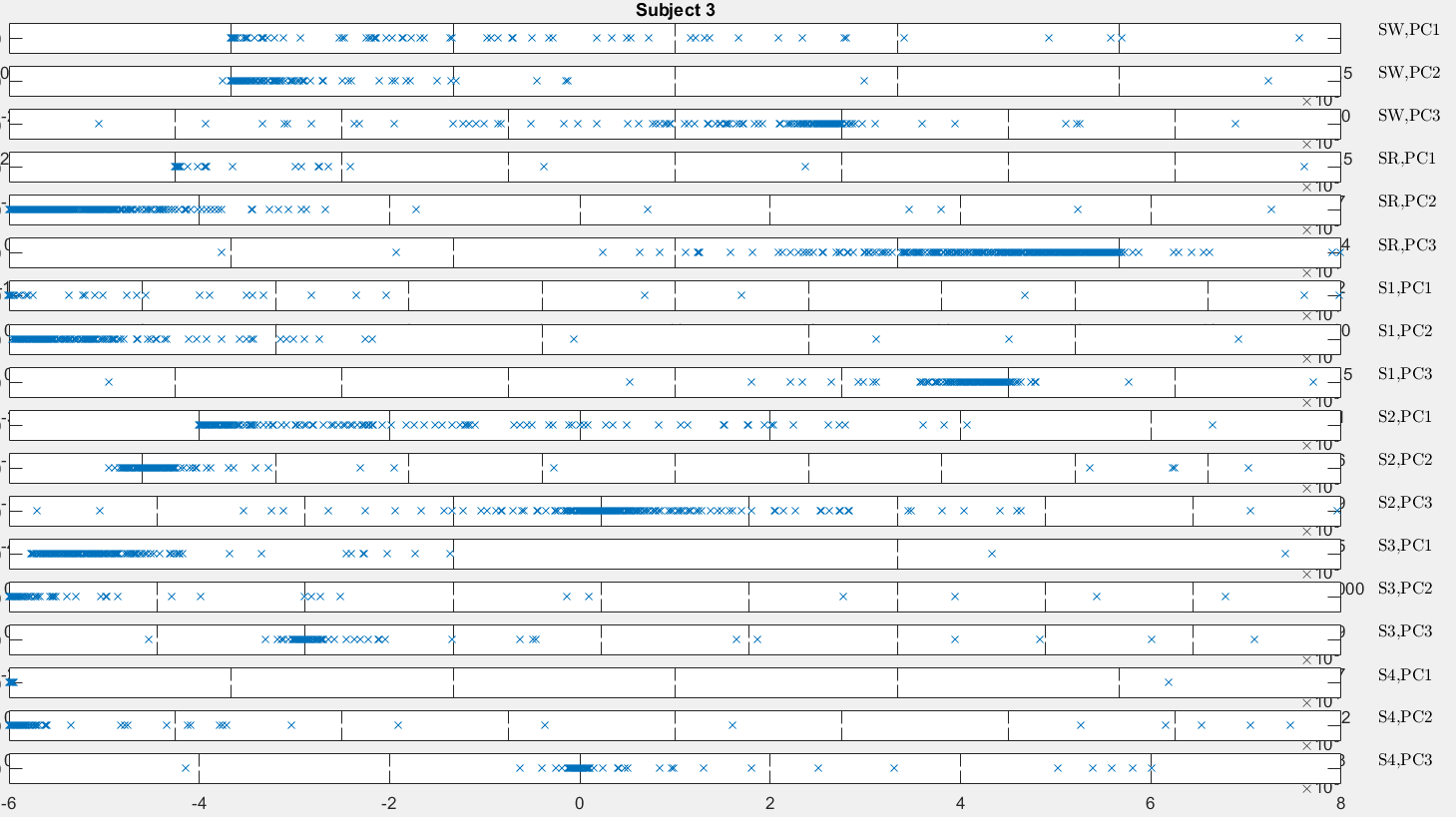
حال، می‎توان نتیجه گرفت که توان متوسط باند فرکانسی دلتا (الکترود Oz) تأثیر بالایی در جداسازی لایه‎های خواب NREM از یکدیگر دارد. همچنین خواب REM را نیز می‎توان تا حدی توسط توان متوسط سیگنال EOG (به شکل دقیق‎تر، ترکیب خطی EOG و EMG) از بقیه جدا کرد.

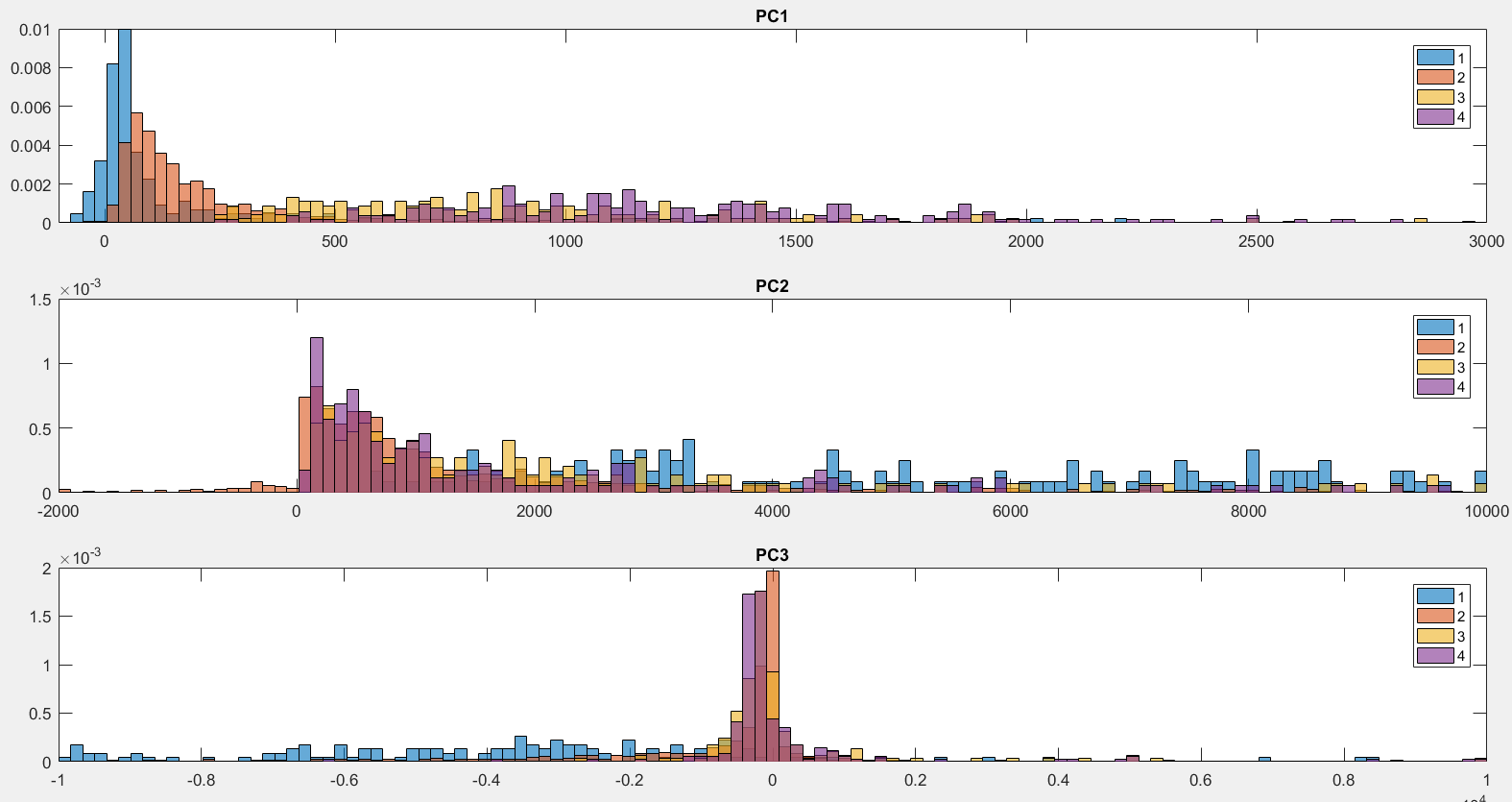
**سابجکت سوم)**





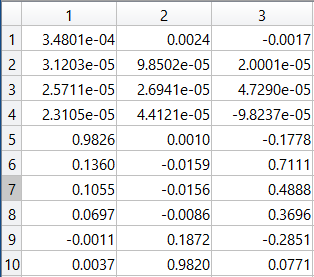






این سابجکت کمی نسبت به دو مورد قبلی متفاوت است. مشاهده می‎شود که راستای PC1 (بر خلاف دو سابجکت قبلی) در جداسازی خواب NREM موفق عمل کرده است. همچنین خواب REM بیشتر در راستای PC2 متمرکز و قابل جداسازی است.

همچنین سه بردار ویژه اصلی برای این سابجکت به صورت زیر است:



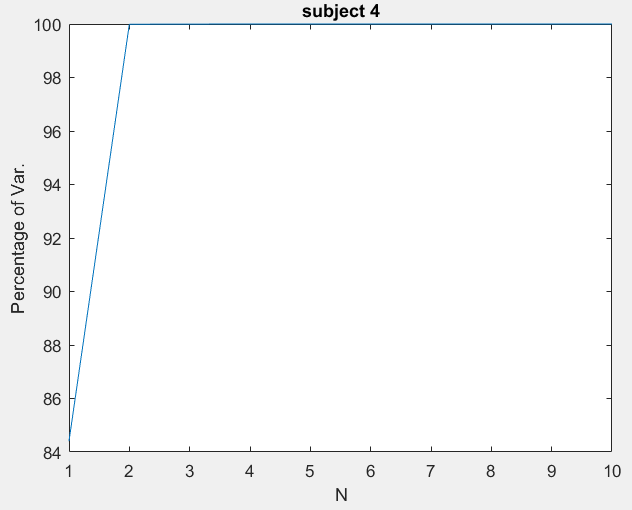
و نتیجه به شرح زیر خواهد بود:

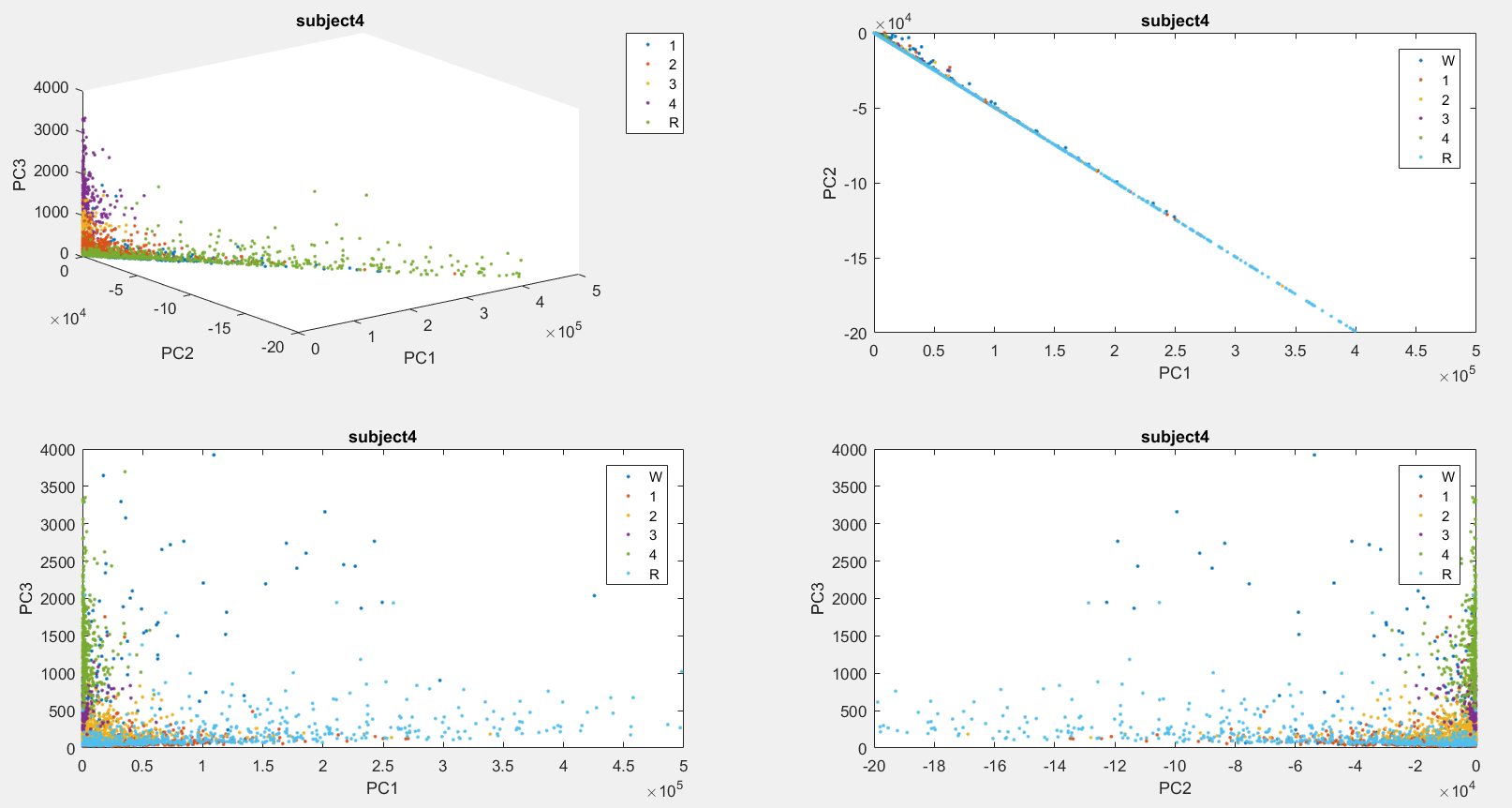
* راستای اول (PC1) : توان متوسط باند فرکانسی دلتای الکترود Oz (سایر پارامترها، اثری ناچیز دارند)
* راستای دوم (PC2) : توان متوسط سیگنال EMG (سایر پارامترها، اثری ناچیز دارند)
* راستای سوم (PC3) : توان متوسط باند فرکانسی تتا الکترود Oz (با ضریب 0.7)، توان متوسط باند فرکانسی آلفای الکترود Oz (با ضریب 0.5)، توان متوسط باند فرکانسی بتای الکترود Oz (با ضریب 0.37)، و توان متوسط سیگنال EOG (با ضریب 0.29-، یعنی اثر معکوس)

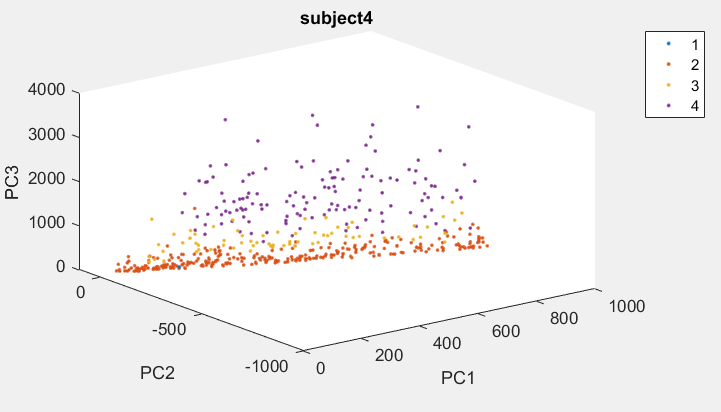
مشاهده می‎شود که توان متوسط باند فرکانسی دلتای الکترود Oz در جداسازی اعماق مختلف خواب NREM و توان متوسط سیگنال EMG در جداسازی خواب REM از سایرین می‎توانند قابل قبول باشند.

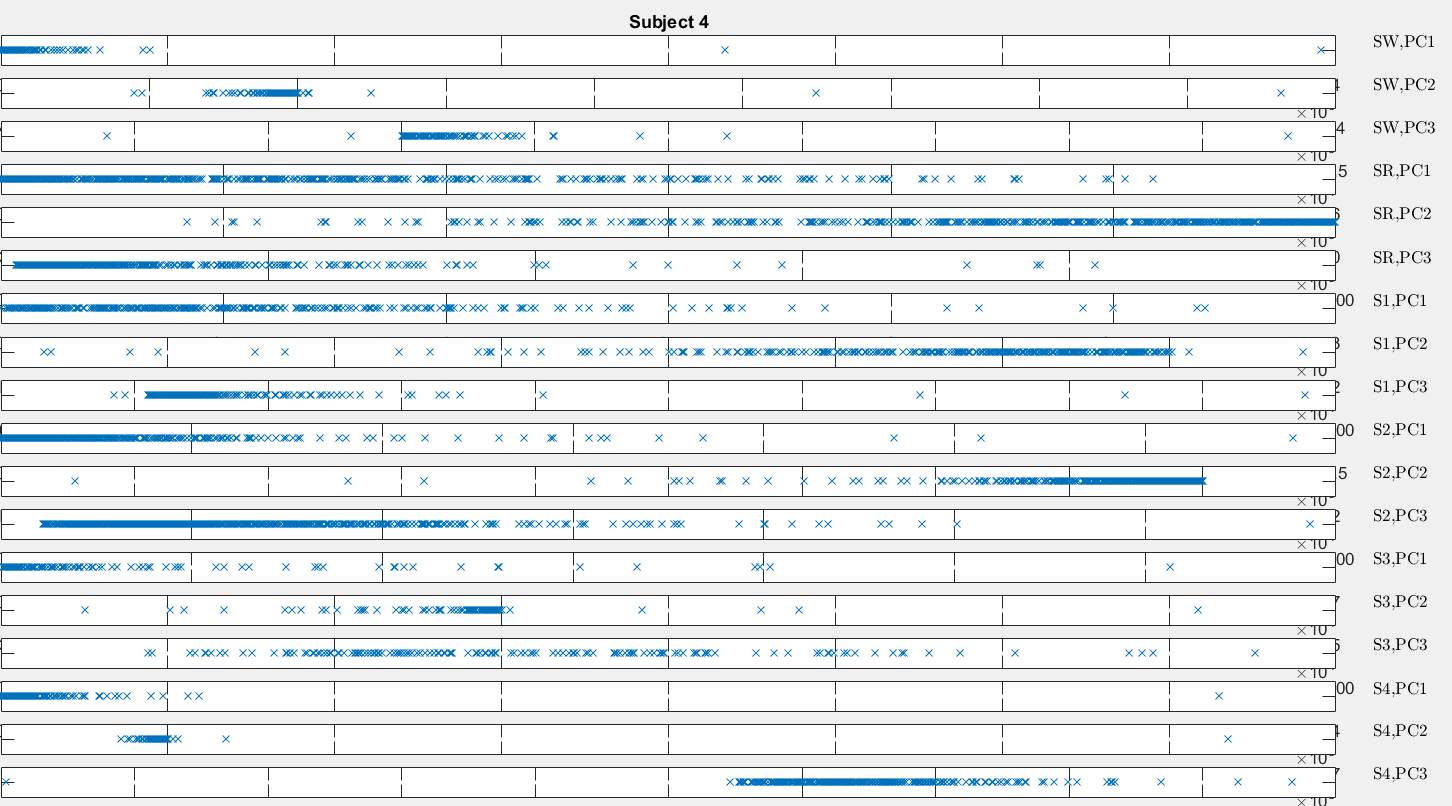
نکته) یک دلیل تفاوت این سابجکت با سایرین، می‎تواند سن وی باشد که 18 سال است، در حالی که سایر سابجکت‎ها در حدود 30 سال دارند.

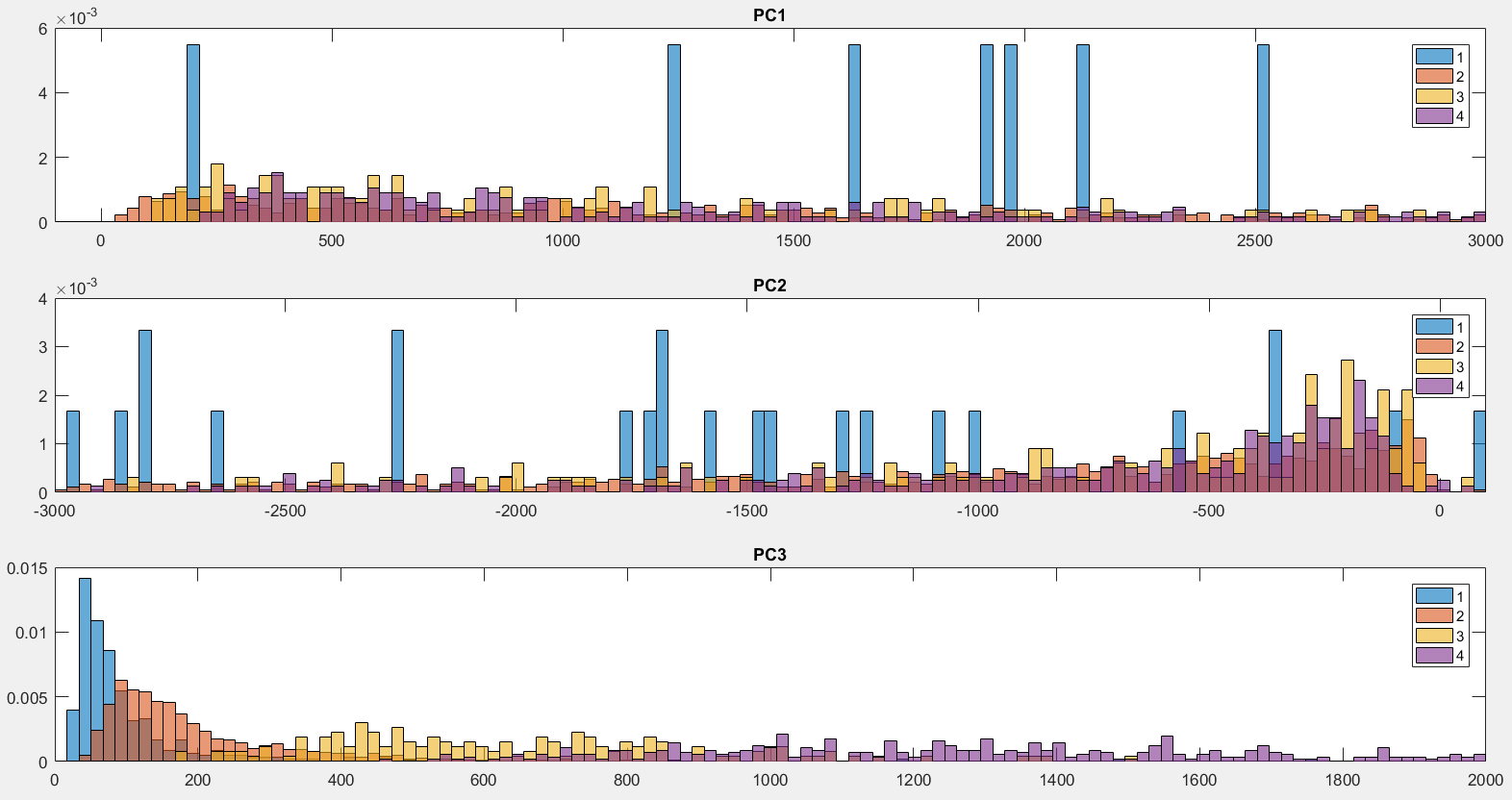
**سابجکت چهارم**







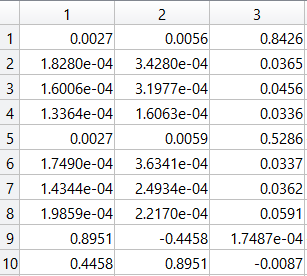




با توجه به شکل‎ها، به خصوص هیستوگرام نهایی، مشاهده می‎شود که راستای PC3 به خوبی حالت‎های NREM را از یکدیگر جدا کرده است.

در خصوص جداسازی حالت REM نیز دو راستای PC1 و PC2 مناسب هستند (چرا که REM در این دو راستا کشیده شده است، در حالی که سایر حالات تقریبا در مبدأ متمرکز هستند)

همچنین سه بردار ویژه اصلی برای این سابجکت به صورت زیر است:

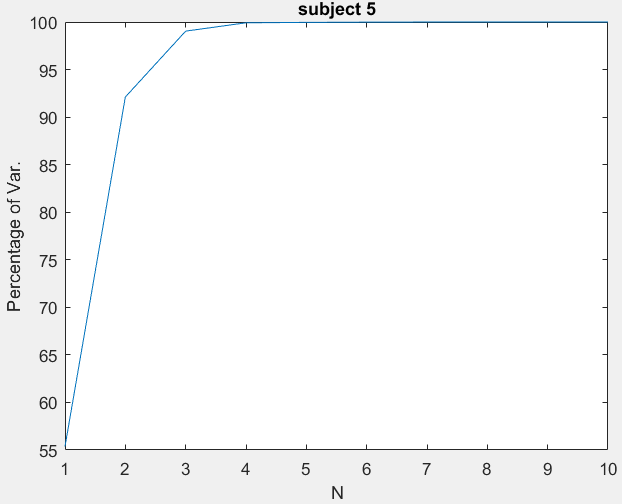


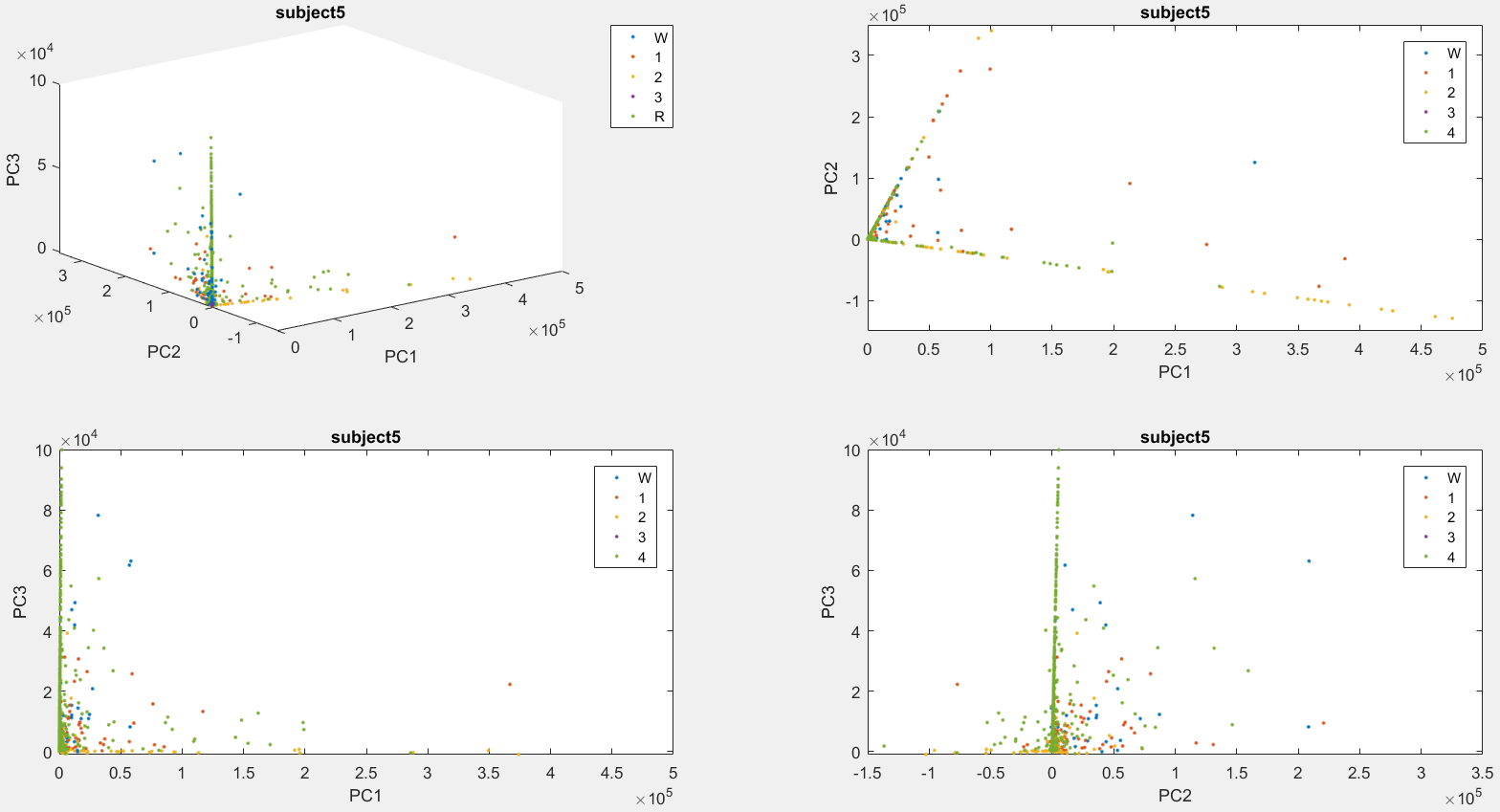
و نتیجه به شرح زیر خواهد بود:

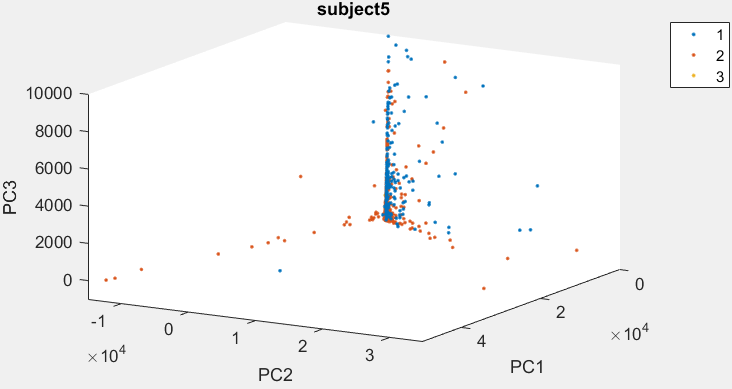
* راستای اول (PC1) : توان متوسط سیگنال EOG (سایر پارامترها، اثری ناچیز دارند)
* راستای دوم (PC2) : توان متوسط سیگنال EMG (سایر پارامترها، اثری ناچیز دارند)
* راستای سوم (PC3) : توان متوسط باند فرکانسی دلتای الکترودهای Fpz (با ضریب 0.85) و Oz (با ضریب 0.53)

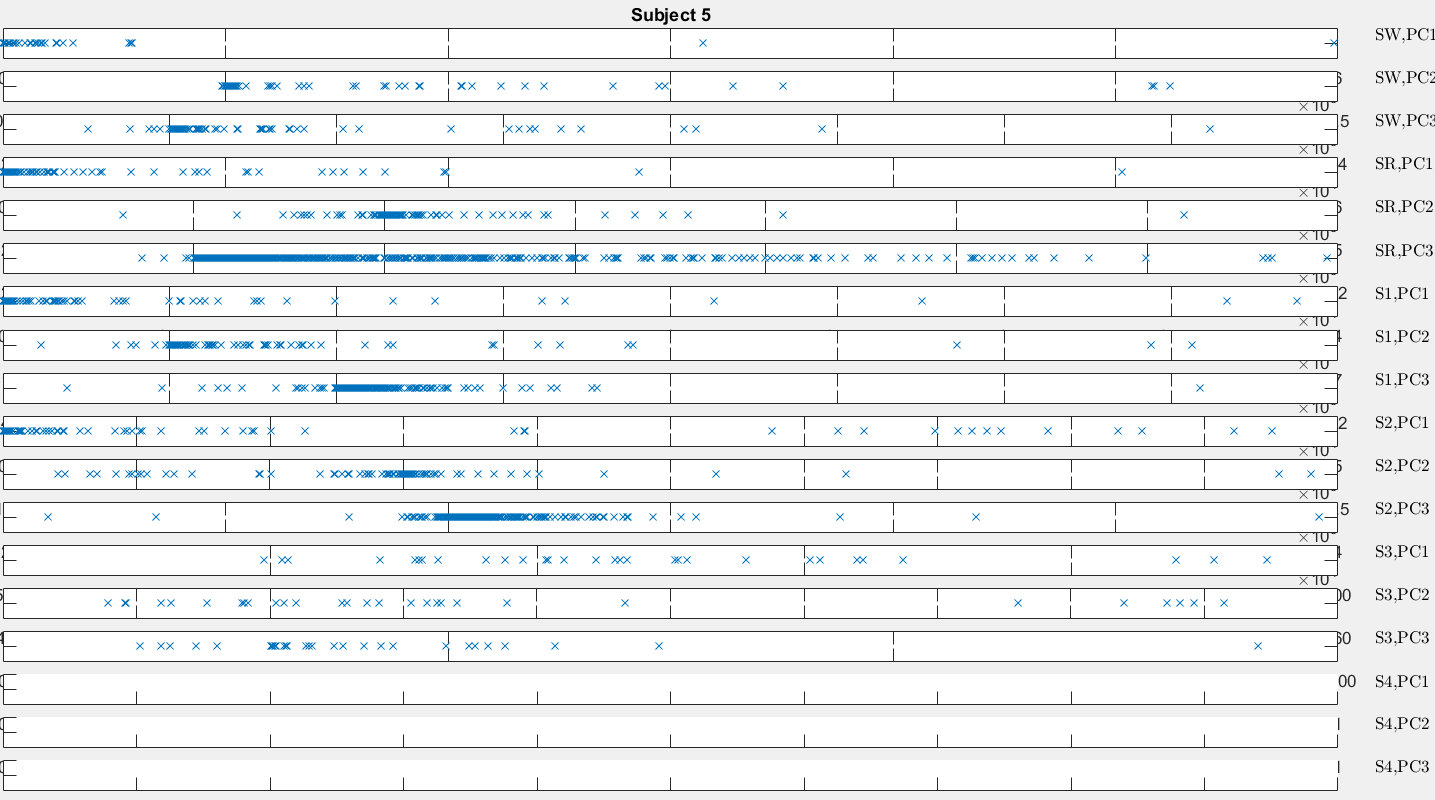
حال، می‎توان نتیجه گرفت که توان متوسط باند فرکانسی دلتا (هم در الکترود Oz و هم در الکترود Fpz) تأثیر بالایی در جداسازی لایه‎های خواب NREM از یکدیگر دارد. همچنین خواب REM را نیز می‎توان تا حدی توسط توان متوسط سیگنالهای EOG و EMG از بقیه جدا کرد.

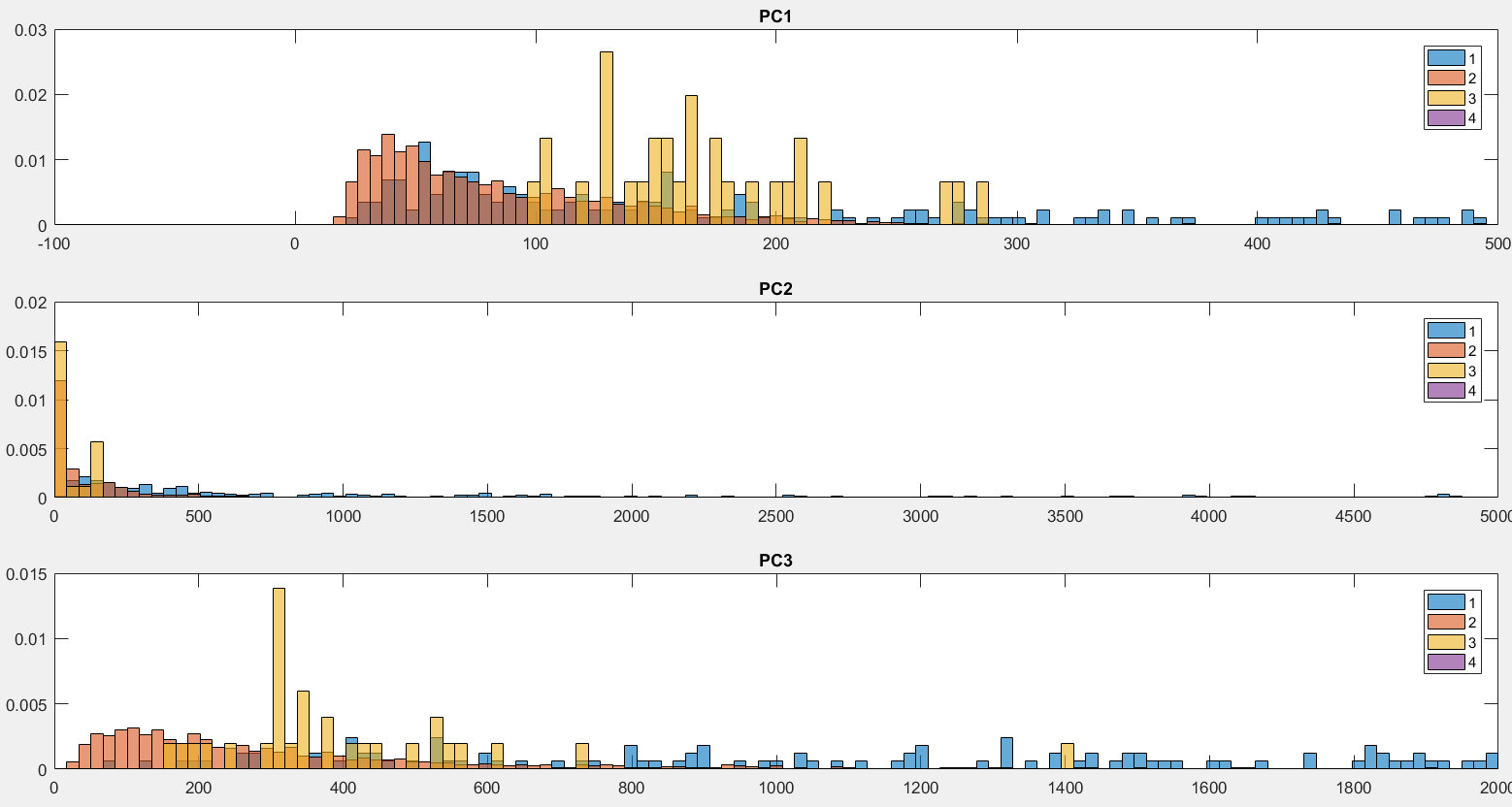
**سابجکت پنجم)**





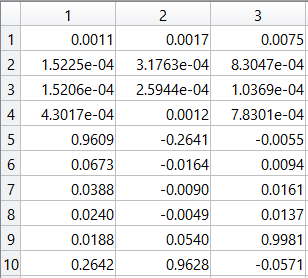






مشاهده می‎شود که (با توجه به هیستوگرام فوق) لایه‎های خواب NREM در راستای PC1 بهترین تفکیک‎پذیری را دارند، اما در مورد خواب REM، در این سابجکت، چیزی که مشاهده می‎شود آن است که خواب REM تقریبا به تمامی در راستای PC3 توزیع شده است و می‎توان به کمک این راستا، آن را از خواب NREM جدا کرد.

همچنین سه بردار ویژه اصلی برای این سابجکت به صورت زیر است:



و نتیجه به شرح زیر خواهد بود:

* راستای اول (PC1) : توان متوسط باند فرکانسی دلتای الکترود Oz (با ضریب 0.96) و توان متوسط سیگنال EMG (با ضریب 0.26)
* راستای دوم (PC2) : توان متوسط سیگنال EMG (سایر پارامترها، اثری ناچیز دارند)
* راستای سوم (PC3) : توان متوسط سیگنال EOG (سایر پارامترها، اثری ناچیز دارند)

حال، می‎توان نتیجه گرفت که توان متوسط باند فرکانسی دلتا (الکترود Oz) تأثیر بالایی در جداسازی لایه‎های خواب NREM از یکدیگر دارد. همچنین خواب REM را نیز می‎توان تا حدی توسط توان متوسط سیگنال EOG از بقیه جدا کرد.

## 3-7)

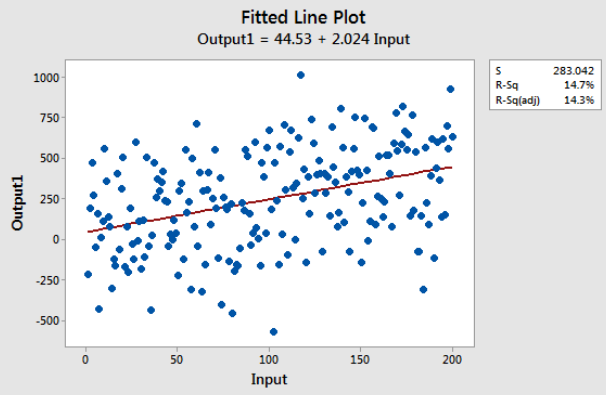
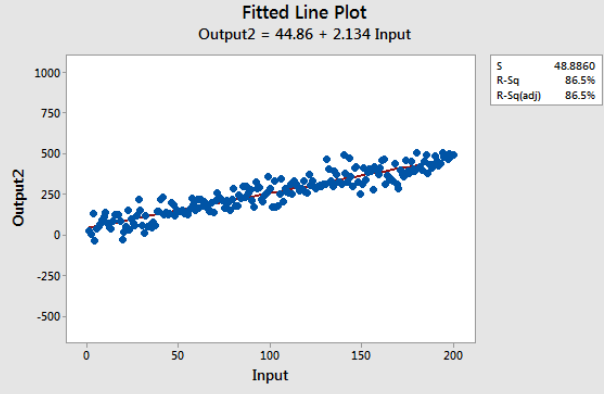
در بخش سوم، یافتن معیاری برای توصیف خواب مبتنی بر PCA، 10 ویژگی (توان متوسط 4 باند فرکانسی الکترود Fpz و Oz و توان متوسط سیگنال‎های EOG و EMG) برای خواب استخراج کردیم و روی آن‎ها PCA زدیم. در نتیجه، 10 راستا در فضای حالت اولیه به دست آوردیم که با توجه به این که PCA ابعاد را بر حسب واریانس داده‎ها مرتب می‎کند، صرفاً سه بُعد اول را مد نظر قرار دادیم و به بررسی این مسأله پرداختیم که هر کدامیک از این راستاها می‎توانند حالت‎های مختلف خواب را از یک‎دیگر جدا کنند. آن‎چه مشاهده شد، تقریباً بین تمامی سابجکت‎های مورد بررسی یکسان بود (نتایج سابجکت‎های مختلف با هم سازگار بودند) و نتیجه این بود که **توان متوسط باند دلتا** (در الکترود Oz) **معیار مناسبی برای جداسازی لایه‎های خواب NREM است** و **توان متوسط سیگنال EOG** (و نیز EMG) **معیار مناسبی برای جداسازی خواب REM از NREM است.**

البته یک نکته قابل ذکر این است که الزاماً برای تمامی سابجکت‎ها، ملاک‎های ذکرشده در جهت‎های یکسانی در فضای 3 بعدی ناشی از PCA نبودند، به عنوان مثال توان متوسط باند دلتا در 3 مورد در راستای PC3 و در 2 مورد در راستای PC1 قرار داشت، لذا نمی‎توان در مورد ارتباط راستاها و لایه‎های خواب ادعای دقیقی کرد، اما در تمامی این موارد، ارتباط ویژگی مربوطه در فضای حالت اولیه با لایه‎های خواب مشترک و سازگار و مطابق توضیحات فوق بودند.

# قسمت چهارم) یافتن معیاری برای عمق خواب مبتنی بر Linear Regression

## 4-1)

در این بخش می‎خواهیم با فیت کردن یک مدل خطی، اثر متغیرها را بر عمق خواب بسنجیم. ابتدا به توضیح دو مشخصه در هر مدل خطی می‎پردازیم:

* **p-value:** در هر مدل خطی، برای هر یک از متغیرهای ورودی، یک p-value محاسبه می‎شود که یک مؤلفه‌ی آماری است که از آزمون فرضیه‎ای با فرض اولیه‎ی **«بی‎اثر بودن این متغیر بر روی پاسخ»** یا به عبارت دیگر، **«صفر بودن ضریب این متغیر»** ناشی می‎شود. بنابراین، هر چقدر p-value کمتر باشد، احتمال بی‎اثر بودن متغیر متناظر کمتر است و این یعنی این متغیر و ضریب به دست آمده برای آن، قابل اعتماد هستند. در حالت کلی، می‎توانیم از بازه اطمینان 95% استفاده کنیم و متغیرهایی با p-valueهای کمتر از 0.05 را بپذیریم.
* **R-squared:** برای هر مدل خطی فیت شده، R-squared یک عدد است که مشخص می‎کند این مدل، تا چه میزان دیتای اصلی را خوب توصیف می‎کند و با کیفیت بالاتری بر روی آن فیت شده است. به عنوان نمونه دو نمودار زیر را در نظر بگیرید:  در هر دو دیتاست، یک خط فیت شده است، اما مشخصاً در یکی از آن‎ها، این خط توصیف‎کننده‎ی بسیار بهتری از دیتای اصلی است، در مقایسه با دیگری. مشاهده می‎شود که در نمونه‎ی سمت راست، مقدار R-squared برابر 0.86 و برای نمونه‌ی سمت چپ، برابر 0.14 است. اگر R-squared در یک مدل خطی برابر 1 باشد، یعنی مدل خطی کاملاً ارتباط بین متغیرها و پاسخ را توصیف می‎کند.

حال می‎خواهیم به بررسی این موضوع بپردازیم که در یک مدل خطی، کدامیک از متغیرها اثر بیشتری (چه اثر مثبت، چه اثر منفی) روی پاسخ دارد و کدام متغیرها تقریباً بی‎اثرند. ابتدایی‎ترین چیزی که به ذهن می‎رسد، آن است که ضریب هر متغیر را بررسی کنیم و هر کدام که اندازه‎ی بزرگ‎تری داشت، اثر بیشتری دارد. ابتدا ببینیم که هر ضریب به چه معنی است: ضریب حاصل از یک مدل خطی برای هر متغیر به معنی «**تغییرات پاسخ، به ازای یک واحد از تغییر هر متغیر مشخص، در حالی که سایر متغیرها ثابت هستند»** می‏باشد. این ادعا که با مقایسه ضریب متغیرها می‎توانیم اهمیت آن‌ها را دریابیم، در حالت کلی غلط است. به عنوان مثال، فرض کنید در یک مدل خطی، یکی از پارامترهای ورودی وزن افراد است. چنان‎چه دیتای مربوط به وزن بر حسب کیلوگرم یا بر حسب گرم بیان شود، بدیهی است اثر وزن در پاسخ تغییر نخواهد کرد، اما ضریب مربوط به وزن، در این دو حالت، از مرتبه 1000 با هم تفاوت خواهند داشت.

برای حل این مشکل، یک روش رایج آن است که دیتای هر کدام از متغیرهای ورودی به صورت استاندارد شده (z-score) ارائه شود.

البته به این روش نیز ایراداتی وارد است، و گفته می‎شود که **نمی‎توان بدون در نظر گرفتن ماهیت مسأله و پارامترهای آن، به یک پاسخ مشخص رسید**، چرا که ممکن است یکی از متغیرها، جدا از ضریب و اطلاعات آماری، قابلیت تغییر چندانی نداشته باشد (مثلا غلظت یک ماد شیمیایی در خون) و لذا برای کنترل یک پدیده، نمی‎توان از این متغیر، هرچند اثرگذاری آن از نظر فرمولی بالا باشد، استفاده کرد. البته این ایراد در این جا به کار ما وارد نیست، چرا که هدف کار ما تغییر دادن عمق خواب با استفاده از تغییر دادن توان متوسط سیگنال‎های ورودی نیست، و ما صرفاً می‎خواهیم مشاهدات خود را تحلیل کنیم و از روی آن استنتاجی به دست آوریم.

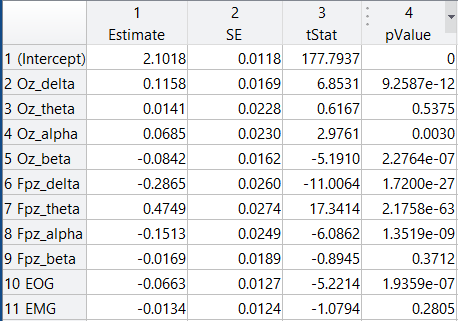
در این جا یک روش دیگر را نیز برای مشکلی که در رابطه با تفاوت بُعد دیتای ورودی مطرح شد، بیان می‎کنیم. روش دوم، بررسی تغییرات R-squared در اثر اضافه شدن یک متغیر، در حالتی است که سایر متغیرها ثابت هستند. این روش نیز روش رایجی است که در آن، مهم‎ترین متغیر را متغیری در نظر می‎گیرند که با اضافه کردن آن، **بیشترین تغییرات مثبت را در R-squared** مشاهده کنیم.

در این جا به این بحث پایان می‎دهیم و در انجام تحلیل خود، از روش اول (محاسبه‎ی z-score و استفاده از دیتای استاندارد شده) استفاده می‎کنیم. البته داده‎های ما مشکل تفاوت ابعادی ندارند (همگی توان متوسط هستند) اما چون مرتبه بزرگی برخی سیگنال‎ها با سایرین متفاوت است، این کار می‎تواند مفید باشد.

مراجع این بخش:

1. <http://blog.minitab.com/blog/adventures-in-statistics-2/how-to-interpret-regression-analysis-results-p-values-and-coefficients>
2. <http://blog.minitab.com/blog/adventures-in-statistics-2/how-to-interpret-a-regression-model-with-low-r-squared-and-low-p-values>
3. <http://blog.minitab.com/blog/adventures-in-statistics-2/how-to-identify-the-most-important-predictor-variables-in-regression-models>
4. <http://blog.minitab.com/blog/adventures-in-statistics-2/how-high-should-r-squared-be-in-regression-analysis>

**نتیجه‎ی سابجکت اول:**



* موثر ترین متغیرها به ترتیب:
  + باند تتای الکترود Fpz (اثر مثبت)
  + باند دلتای الکترود Fpz (اثر منفی)
  + باند آلفای الکترود Fpz (اثر منفی)
  + باند دلتای الکترود Oz (اثر مثبت)
  + مابقی متغیرها تقریباً بی‎اثرند

همچنین چهار متغیری که نام برده شدند، p-valueهای بسیار ناچیزی دارند که همان‎طور که توضیح داده شد، اثر آن‎ها می‎تواند از نظر آماری قابل اعتنا باشد.

## 4-2)

توضیحات کامل در مورد مفهوم R-squared در بخش قبلی بیان شد.

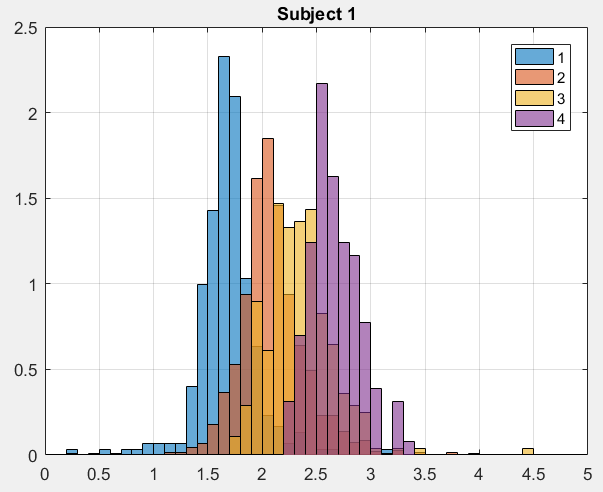
برای این سابجکت، مقدار R-squared برابر با 0.32 است. با توجه به توضیحاتی که در بخش قبل داده شد، می‎توان گفت که مدل توصیف‎کننده‎ی چندان جالبی نیست، چرا که مقادیر R-squared بین صفر و یک تغییر می‎کند و مقدار 0.32 نسبتاً پایین است. البته این سؤال که **«R-squared خوب چقدر است؟»** سؤال درستی نیست، چرا که بسته به کاربرد ما از مدل، ممکن است مقادیر مختلفی از این پاسخ برای ما کافی باشد، اما در حالت کلی می‎توان گفت که مدل ما، **32% از تغییرات پاسخ، توسط متغیرها توصیف می‎شود**.

## 4-3)

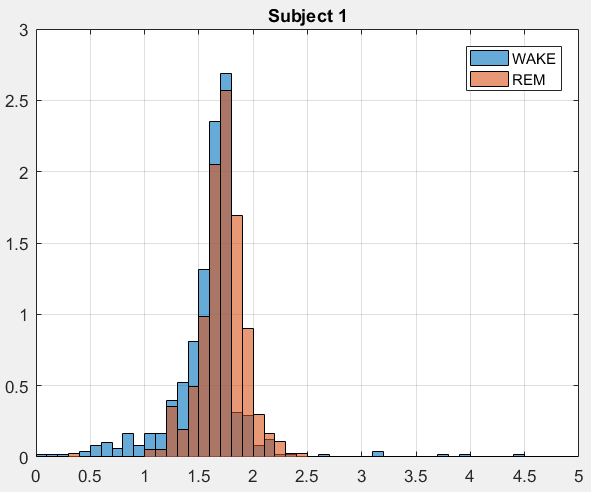
آن چه تا به این جا به دست آورده‎ایم، بیشتر یک مدل **توصیف‎کننده** است تا پیش‎بینی‎کننده. علت این امر آن است که برای ما، آن‎چه مهم‎تر بوده است، آن است که کدام متغیرها دیتای موجود را بهتر توصیف می‎کنند و اثر مشخص‎تری دارند (سؤال اول) و نیز آن که این مدل، در کل، چه میزان از تغییرات پاسخ (عمق خواب) را به درستی توسط متغیرهای موجود توصیف می‎کند. (سؤال دوم)

در واقع، ما تا به این جا سعی کرده‎ایم پارامترهای مؤثر را شناسایی کنیم تا بتوانیم ملاکی برای عمق خواب بیابیم، که این کار بیشتر در محدوده‎ی توصیف قرار می‎گیرد و نه پیش‎بینی. (البته این کار، می‎تواند مقدمه‎ای باشد برای انجام یک فعالیت پیش‎بینی‎کننده)

## 4-4)



## 4-5)

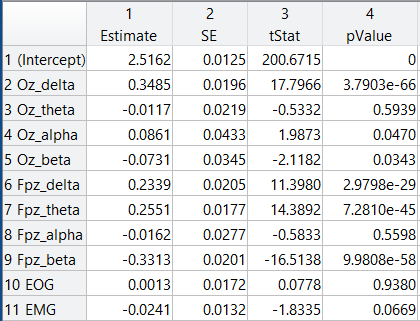


مشاهده می‎شود که هر دو وضعیت خواب REM و بیداری، عمق پیش‎بینی شده‎ای دارند که اکثریت تمرکز آن در محدود 1 تا 2 است. این پیش‎بینی به شکل حدودی منطقی است، چرا که باید از عمق‎های بالای خواب فاصله‎ی زیادی داشته باشیم که در این جا نیز چنین مشاهده می‎شود. به شکل دقیق‎تر، انتظار داریم که عمق پیش‎بینی‎شده برای بیداری، کمترین مقدار ممکن را به خود گرفته، و الزاماً کمتر از 1 باشد، که این انتظار در این جا برآورده نشده، البته می‎بینیم که توزیع بیداری، در مقایسه با توزیع REM بیشتر به سمت عمق‎های کم گرایش دارد که این نتیجه‎ی درستی است.

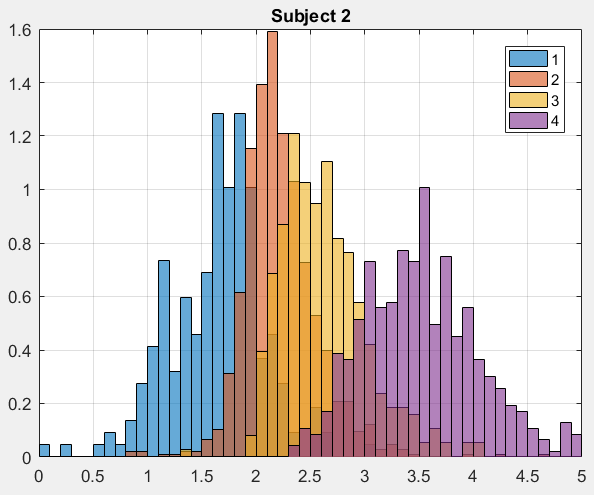
البته طبعاً انتظار نداریم که این مدل خطی همه چیز را به درستی توصیف کند، و ضمناً با بررسی سابجکت‎های دیگر در بخش‎های بعدی، به جمع‎بندی بهتری خواهیم رسید.

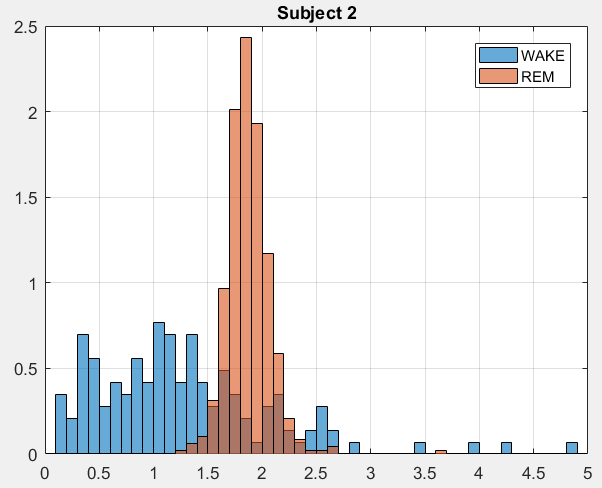
## 4-6)

**سابجکت دوم:**



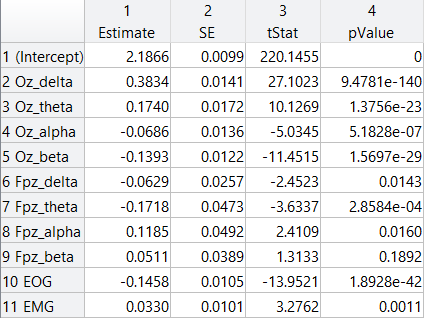
* موثر ترین متغیرها به ترتیب:
  + باند دلتای الکترود Oz (اثر مثبت)
  + باند بتای الکترود Fpz (اثر منفی)
  + باند تتای الکترود Fpz (اثر مثبت)
  + باند دلتای الکترود Fpz (اثر مثبت)
  + مابقی متغیرها تقریباً بی‎اثرند
* مقدار R-squared برابر با 0.63 است، و مدل برای این سابجکت توصیف نسبتاً خوبی دارد، به طوری که 63 درصد تغییرات حالت خواب توسط متغیرها توصیف می‎شود.



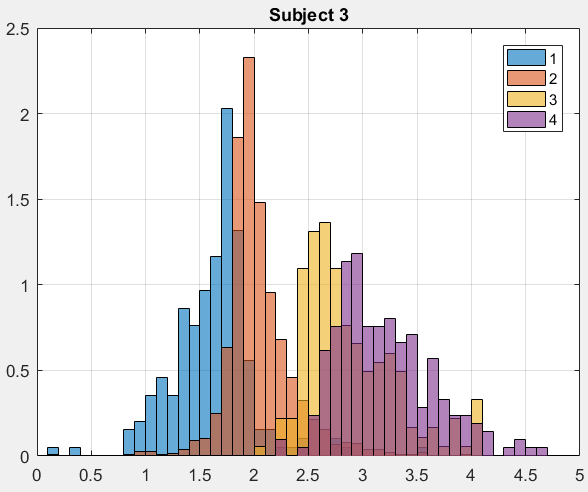


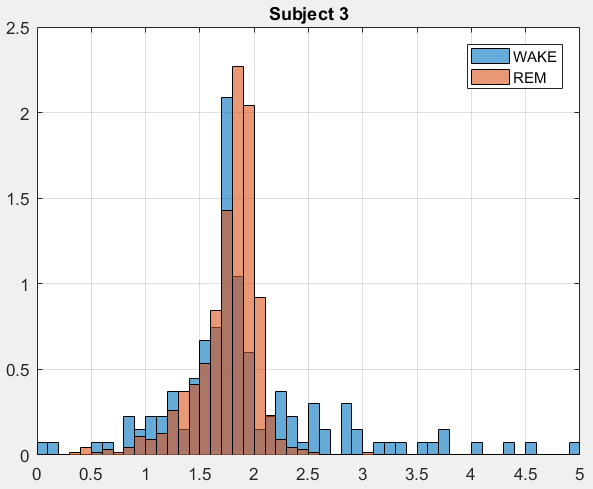
میزان جداشدگی حالت بیداری و REM و عمق پیش‎بینی‎شده برای آن‎ها در این سابجکت، خوب و منطقی است، به طوری که حالت بیداری کمترین عمق، و به خصوص، عمقی کمتر از 1 را به خود اختصاص می‎دهد، و دیتای REM نیز عمقی در حد فاصل 1 تا 2 دارد که با انتظارات ما، تناسب نسبی دارد، به خصوص که عمق هیچ کدام از این دو وضعیت، به حالت‎های عمیق خواب نزدیک نمی‎شود.

**سابجکت سوم:**



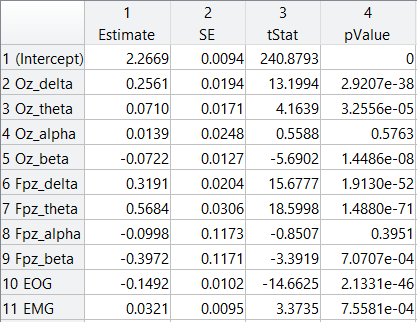
* موثر ترین متغیرها به ترتیب:
  + باند دلتای الکترود Oz (اثر مثبت)
  + باند تتای الکترود Oz (اثر مثبت)
  + باند تتای الکترود Fpz (اثر منفی)
  + توان متوسط سیگنال EOG (اثر منفی)
  + باند بتای الکترود Oz (اثر منفی)
  + باند آلفای الکترود Fpz (اثر مثبت)
  + مابقی متغیرها تقریباً بی‎اثرند
* مقدار R-squared برابر با 0.59 است، و مدل برای این سابجکت توصیف نسبتاً خوبی دارد، به طوری که 59 درصد تغییرات حالت خواب توسط متغیرها توصیف می‎شود.



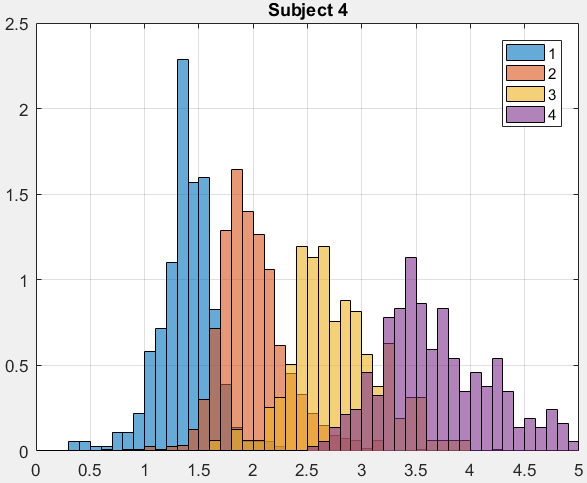


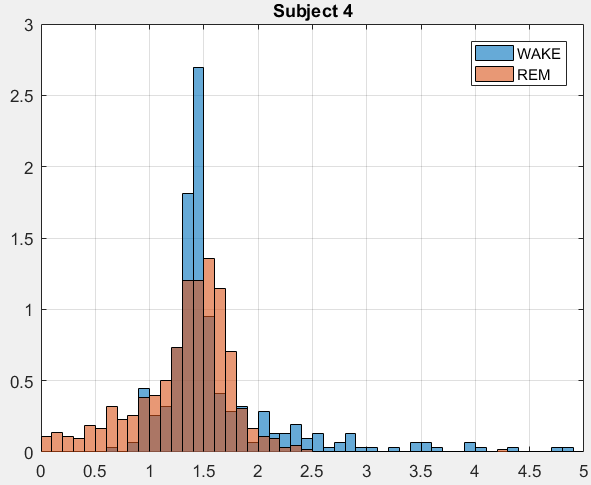
میزان جداشدگی حالت بیداری و REM و عمق پیش‎بینی‎شده برای آن‎ها در این سابجکت، نسبت به موارد قبلی خیلی جالب نیست، به خصوص که بیداری و REM تقریباً بر هم منطبق هستند، و نیز دیتای بیداری حضوری (اگرچه کم) در اعماق بالاتر نیز دارد. با این وجود، تمرکز هر دو دسته داده در حد فاصل اعماق 1 تا 2 است، که این با موارد قبلی (به خصوص در مورد خواب REM) سازگار است.

**سابجکت چهارم:**



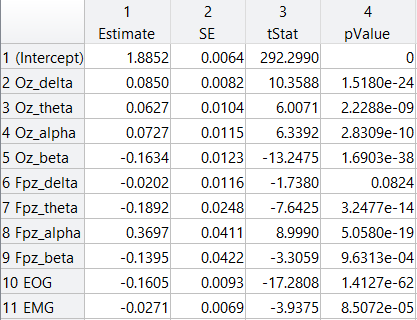
* موثر ترین متغیرها به ترتیب:
  + باند تتای الکترود Fpz (اثر مثبت)
  + باند بتای الکترود Fpz (اثر منفی)
  + باند دلتای الکترود Fpz (اثر مثبت)
  + باند دلتای الکترود Oz (اثر مثبت)
  + توان متوسط سیگنال EOG (اثر منفی)
  + باند آلفای الکترود Fpz (اثر منفی)
  + مابقی متغیرها تقریباً بی‎اثرند
* مقدار R-squared برابر با 0.8 است، و مدل برای این سابجکت توصیف خوبی دارد، به طوری که 80 درصد تغییرات حالت خواب توسط متغیرها توصیف می‎شود.



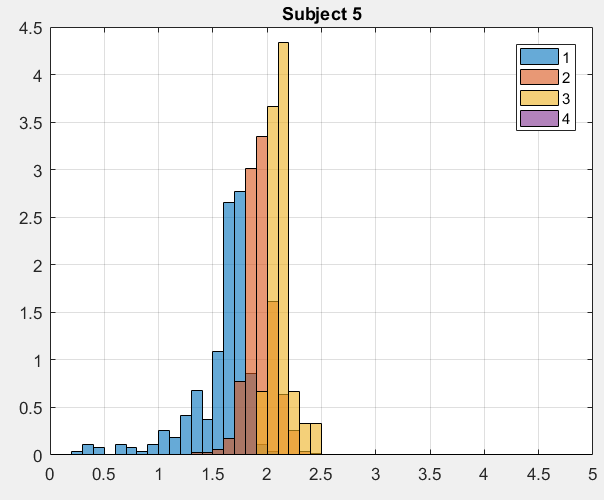


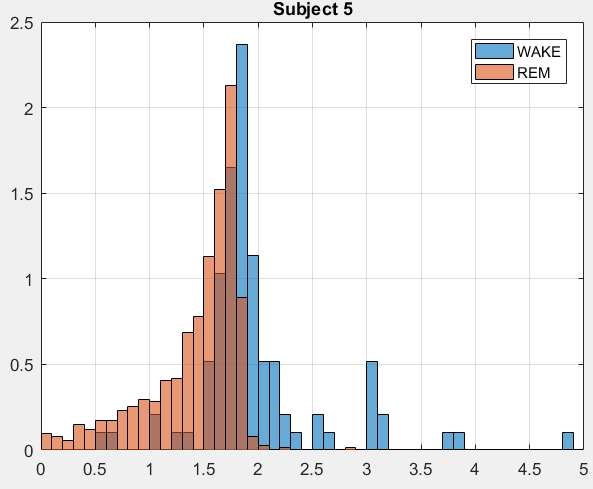
مجدداً در این حالت نیز تمرکز عمق پیش‎بینی‎شده، هم برای REM، و هم برای بیداری، در محدوده 1 تا 2 است، با این تفاوت که این بار، کشیدگی REM به سمت اعماق کمتر و کشیدگی بیداری به سمت اعماق بیشتر است، که این نتیجه نادرست است و نشان می‎دهد که پیش‎بینی انجام شده، ایراد دارد.

**سابجکت پنجم:**



* موثر ترین متغیرها به ترتیب:
  + باند آلفای الکترود Fpz (اثر مثبت)
  + باند تتای الکترود Fpz (اثر منفی)
  + باند بتای الکترود Oz (اثر منفی)
  + توان متوسط سیگنال EOG (اثر منفی)
  + باند بتای الکترود Fpz (اثر منفی)
  + مابقی متغیرها تقریباً بی‎اثرند
* مقدار R-squared برابر با 0.35 است، و مدل برای این سابجکت توصیف چندان جالبی ندارد، به طوری که 35 درصد تغییرات حالت خواب توسط متغیرها توصیف می‎شود.





مجدداً در این حالت نیز تمرکز عمق پیش‎بینی‎شده، هم برای REM، و هم برای بیداری، در محدوده 1 تا 2 است، با این تفاوت که این بار، کشیدگی REM به سمت اعماق کمتر و کشیدگی بیداری به سمت اعماق بیشتر است، که این نتیجه نادرست است و نشان می‎دهد که پیش‎بینی انجام شده، ایراد دارد.

**جمع‎بندی:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | سابجکت اول | سابجکت دوم | سابجکت سوم | سابجکت چهارم | سابجکت پنجم |
| متغیر اول |  |  |  |  |  |
| متغیر دوم |  |  |  |  |  |
| متغیر سوم |  |  |  |  |  |
| متغیر چهارم |  |  |  |  |  |
| متغیر پنجم |  |  |  |  |  |
| متغیر ششم |  |  |  |  |  |

جدول فوق، نتایج بخش اول تحلیل هر سابجکت (کدام متغیرها اثر بیشتری بر عمق خواب دارند) را به شکل خلاصه‎شده نمایش می‎دهد.

باند تتای الکترود Fpz در تمامی ستون‎های این جدول حضور دارد، و این یعنی به نظر، یکی از مهم‎ترین مؤلفه‎های تشخیص عمق خواب است. البته در دو مورد اثر منفی و در سه مورد اثر مثبت دارد که این کمی نتیجه‎گیری کلی ما را تحت تأثیر قرار می‎دهد، چرا که نمی‎توان به تنهایی، رابطه‎ی مشخصی بین انرژی باند تتای الکترود Fpz و عمق خواب تعیین کرد.

باند دلتای الکترود Oz و باند آلفای الکترود Fpz در چهار ستون از پنج ستون جدول فوق حضور دارند. باند آلفای Fpz در سه مورد اثر منفی و در یک مورد اثر مثبت داشته است، و باند دلتای الکترود Oz در تمامی موارد اثر مثبت داشته است. همچنین در یک موردی هم که این متغیر به عنوان متغیرهای تأثیرگذار شناسایی نشد، باز هم اثر آن مثبت (هر چند ناچیز) بوده است.

در ادامه، باند دلتای الکترود Fpz، باند بتای الکترود Fpz و سیگنال EOG نیز سه حضور در جدول فوق داشته اند.

در جدول فوق، 25 متغیر در 5 ستون نام برده شده اند که سه متغیر فوق (باند تتای الکترود Fpz، باند دلتای الکترود Oz و باند آلفای الکترود Fpz)، در 13 مورد از این 25 مورد مشاهده می‎شوند، یعنی بیش از 50 درصد موارد را این سه به خود اختصاص داده‎اند و می‎توان این سه را به عنوان پارامترهای برتر در نظر گرفت.

بنابراین به عنوان جمع‎بندی می‎توان گفت که این مشاهدات برای سابجکت‎های مختلف یکسان نیست، اما می‎توان ویژگی‎هایی را معرفی کرد که تقریبا در تمامی سابجکت‎ها اثرگذار بوده‎اند و می‎توان ادعا کرد که این ویژگی‎ها، "ویژگی‎های مؤثر" هستند.

در مورد بخش آخر، یعنی پیش‎بینی عمق خواب REM و حالت بیداری نیز، با توجه به مشاهدات انجام شده، باید گفت که این کار چندان دقیق نیست، اما نتایج در حالت کلی عمدتاً در محدوده عمق 1 تا 2 توزیع شده‎اند، که از نظر فاصله داشتن با اعماق بالا قابل قبول است، اما به خصوص برای بیداری انتظار داریم که عمق بسیار پایین و به ویژه کمتر از 1 داشته باشد، که به جز یک سابجکت، سایرین چنین نتیجه‎ای ارائه ندادند و این یعنی نمی‎توان از این برون‎یابی، انتظار چندانی داشت. همچنین یک نکته دیگر آن است که در اکثر مشاهدات، تمایز چندانی بین حالت بیداری و REM مشاهده نمی‎شد.

## 4-6)

بخشی از استدلال‎های این بخش، در صفحه‎ی قبل بیان شد. با استناد به توضیحات صفحه قبل و جدول رسم شده، به این نتیجه می‎رسیم که پارامترهای تأثیرگذار برای سابجکت‎های مختلف به طور دقیق یکسان نیستند، به خصوص که ترتیب اثرگذاری آن‎ها دقیقاً بر هم منطبق نیست، اما با بررسی پارامترهای مؤثر، می‎بینیم که سه پارامتر وجود دارند که تقریباً در تمامی سابجکت‎ها اثرگذار هستند. این سه عبارتند از باند تتای الکترود Fpz، باند دلتای الکترود Oz، و باند بتای الکترود Fpz.

از میان این سه، **باند دلتای الکترود Oz** تنها موردی است که در تمامی موارد، اثر آن از نظر مستقیم یا معکوس بودن یکسان است، و برای تمامی پنج سابجکت مورد بررسی، اثر مثبتی بر عمق خواب داشته است، یعنی با افزایش مقدار آن، عمق خواب نیز زیاد شده است.

این نتیجه، دقیقاً همان نتیجه‎ای است که در قسمت تحلیل PCA نیز به دست آمده بود. (و نیز مقالات علمی منتشرشده نیز تا حد قابل قبولی آن را تأیید می‎کنند)

در حالت کلی نیز باید توجه داشت که این دو تحلیل (PCA و Linear Regression) ماهیت‎هایی متفاوت و فرض‎هایی متفاوت دارند و نباید انتظار داشت که نتایج آن‎ها کاملاً منطبق باشند.

ضمناً در تحلیل PCA، توان متوسط سیگنال EOG به عنوان پارامتر جداکننده خواب REM شناسایی شد که در این بخش، ما به طور کلی بخش مربوط به خواب REM را از دیتاست حذف کردیم و انتظار نداریم نتیجه‎ی متناظری به دست بیاید.

در مورد R-squaredها، به طور متوسط مقادیری بیش از 50% مشاهده می‎شد که می‎توانست قابل قبول باشد.

نهایتاً در مورد پیش‎بینی عمق برای بیداری و REM، دیدیم که نتایج چندان جالب توجه نیستند و انتظاراتی از این قبیل که REM و بیداری جدا شوند، یا این که عمق بیداری تفاوت معنی‎داری با صفر نداشته باشد را برآورده نمی‎کنند، و عمق پیش‎بینی‎شده برای هر دو، در فاصله 1 تا 2 متمرکز است؛ اما از این نظر که با اعماق بالا تفاوت معنی‎داری مشاهده می‎شود، نتایج درست و معقول هستند.

# قسمت پنجم) یک سؤال دلخواه!

در این بخش به بررسی چند سؤال مختلف پرداخته‎ایم.

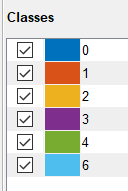
## سؤال اول) می‎خواهیم با استفاده از طبقه‎بندی داده‎های موجود با توجه به ویژگی‎های استخراج شده (ماتریس X)، و با اعتبارسنجی Cross Validation ببینیم که این ویژگی چه میزان موفق به جداسازی حالت‎های مختلف خواب می‎شود.

برای پاسخ‎گویی به این قسمت، از روش طبقه‎بندی خطی SVM استفاده می‎کنیم. برای این کار، یک راه استفاده از زوج‎های توابع متلب مثل svmtrain و svmclassify یا fitcsvm و predict است، که البته این توابع فرایند طبقه‎بندی را برای دو کلاس انجام می‎دهند، لذا باید برای n کلاس، ابتدا n طبقه‎بندی‎کننده‎ی دو کلاسه آموزش دهیم، سپس دیتای تست را به تمامی آن‎ها بدهیم، و هر کدام که دیتا را جزء گروه خود اعلام کرد، جواب مسأله را به ما می‎دهد. (وظیفه هر طبقه‎بندی‎کننده، جداسازی دیتای یک گروه از سایر گروه‎ها است) بعد از این کار، باید با روش اعتبارسنجی Cross Validation، ببینیم که ویژگی‎های استخراج شده، تا چه حد می‎توانند دیتای مربوط به دسته‎های مختلف را از هم جدا کنند. روش k-fold Cross Validation داده‎های آموزش را به k دسته تقسیم می‎‎کند، سپس فرایند طبقه‎بندی را k بار انجام می‎دهد که هر بار، یک دسته به عنوان دیتای تست، و k-1 دسته به عنوان دیتای آموزش به کار می‎روند. نهایتاً درصد پاسخ‎گویی صحیح در این k آزمایش را میانگین می‎گیرد و به عنوان درصد پاسخ‎گویی نهایی اعلام می‎کند. در این مسأله ما داده‎ی جدیدی برای طبقه‎بندی نداریم، لذا صرفاً نتیجه‌ی فرایند Cross Validation است که حائز اهمیت است.

یک نکته در مورد Cross Validation و این که درصد پاسخ آن چقدر قابل اعتماد است وجود دارد، و آن نکته این است که گاهی ویژگی‎های استخراج شده، اثری (مانند یک آماره) از کل دیتای آموزش را در خود دارند، لذا اگرچه در فرایند اعتبارسنجی، داده‎ها به دو دسته تقسیم شده اند و ما به ظاهر، از داده‎های تست کاملا بی‎اطلاعیم، اما ویژگی‎های استخراج شده دارای یک بایاس مرتبط با مجموعه کلی داده‎ها هستند، که می‎تواند باعث شود نتیجه‎ Cross Validation بهتر از نتیجه‎ای شود که در صورت آزمایش روی داده‎های جدید به دست آید. اگر این مشکل وجود نداشته باشد، می‎توان ادعا کرد که این اعتبارسنجی کاملاً مشابه مسأله‎ای است که در آن، دیتای تست جدید و ناشناخته‎ای در اختیار ما قرار داده باشند و بخواهیم آن‎ها را طبقه‎بندی کنیم.

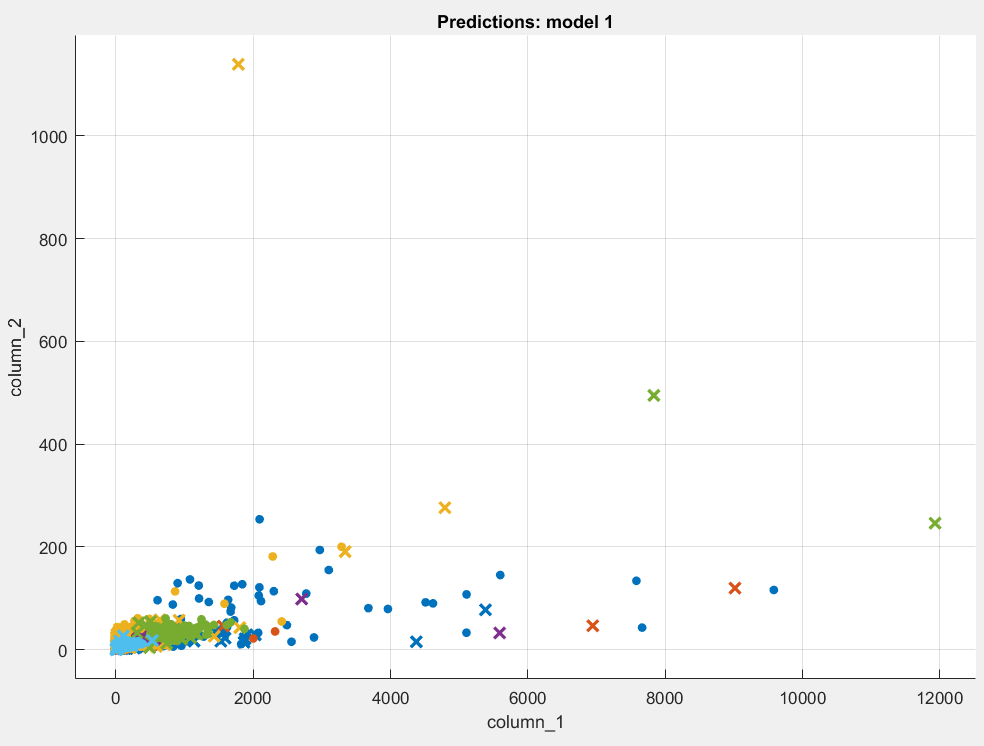
همان طور که اشاره شد، در این جا ما صرفاً می‎خواهیم ببینیم که این ویژگی‎ها چه میزان قادر به جداسازی هستند و صرفاً نتیجه فرایند Cross Validation برای ما اهمیت دارد. برای پیاده‎سازی این فرایند، از ابزار Classification Learner متلب استفاده می‎کنیم و این کار را برای هر 5 سابجکت انجام می‎دهیم و برای هر یک، **نمودار دیتای درست و اشتباه تشخیص داده شده در دو راستا، نتیجه Cross Validation و درصد صحت، و confusion matrix** (نشان‎دهنده این است که تشخیص‎های اشتباه بین کدام دو گروه اتفاق افتاده‎اند) را گزارش می‎کنیم، و در نهایت به جمع‎بندی می‎پردازیم.

ضمنا در تمامی نمودارهای پراکندگی رسم‎شده، راهنمای نمودار به صورت زیر است که شماره کلاس‎ها، متناظر با بردار state است که در طول پروژه با آن کار کرده‎ایم.



**سابجکت اول:**

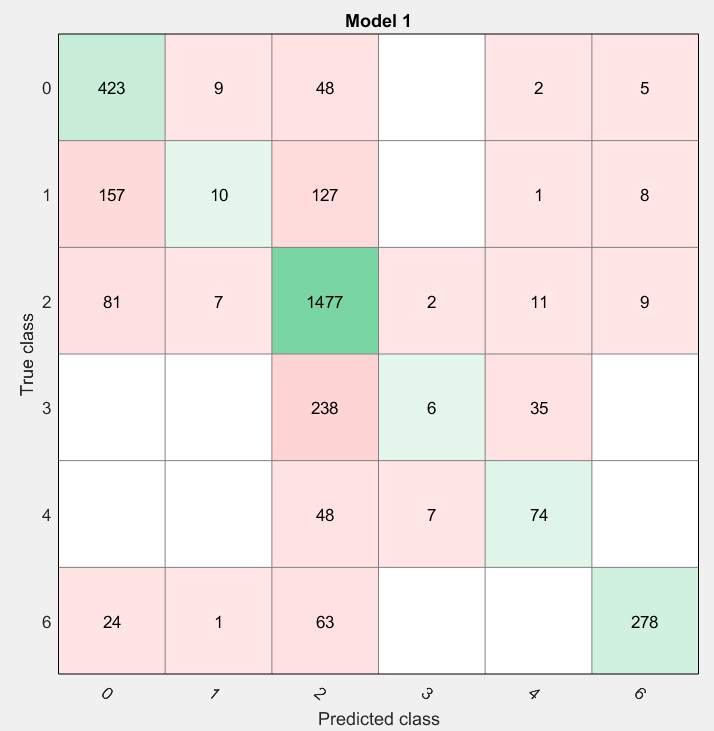
الف) نمودار پراکندگی در دو راستا (ضربدر ها نقاطی هستند که اشتباه تشخیص داده شده اند)



ب) حاصل Cross Validation

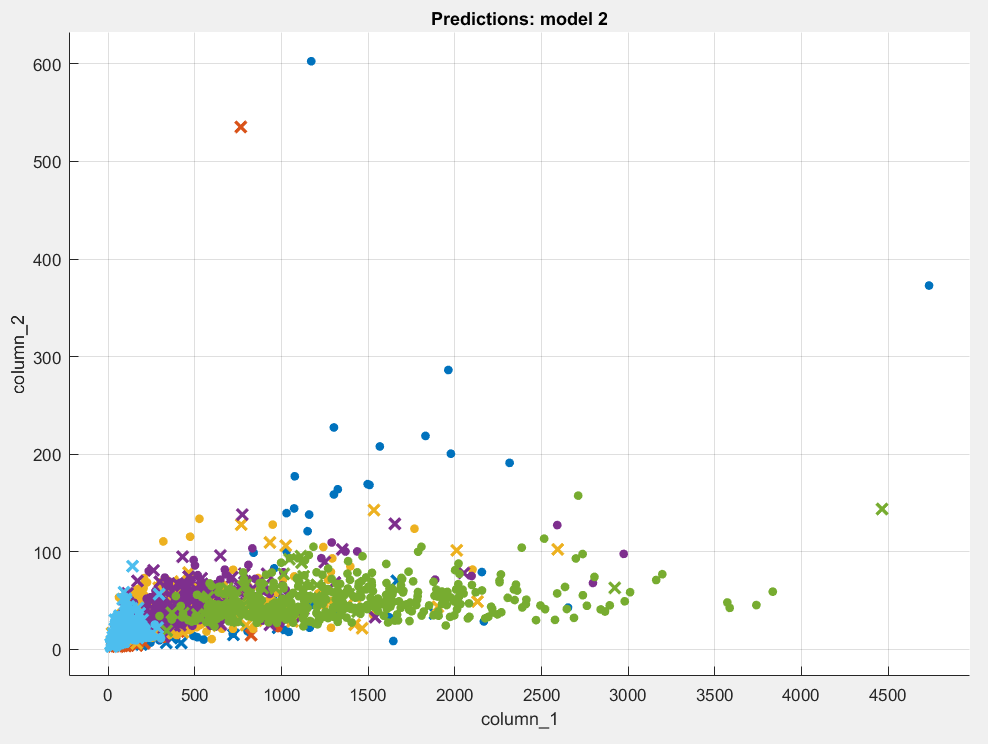


ج) Confusion Matrix



**سابجکت دوم:**

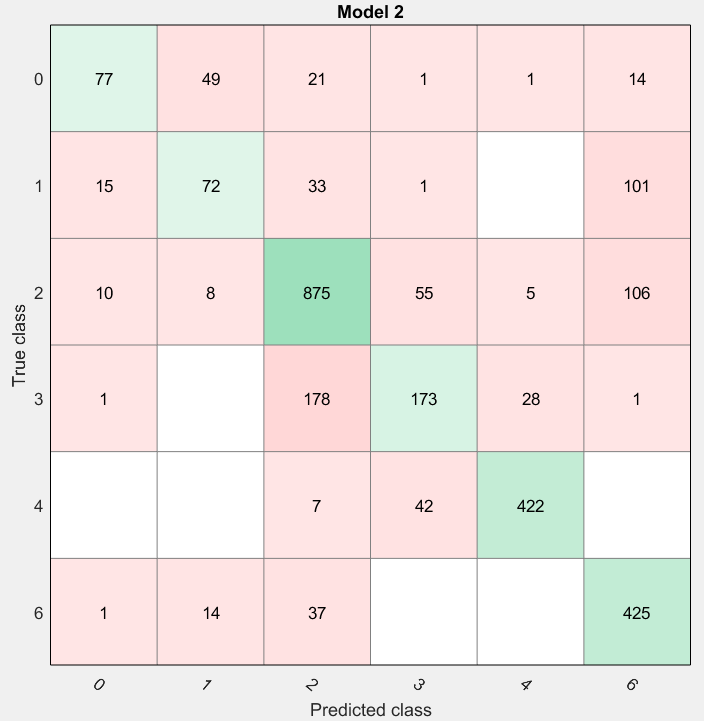
الف) نمودار پراکندگی در دو راستا (ضربدر ها نقاطی هستند که اشتباه تشخیص داده شده اند)



ب) حاصل Cross Validation

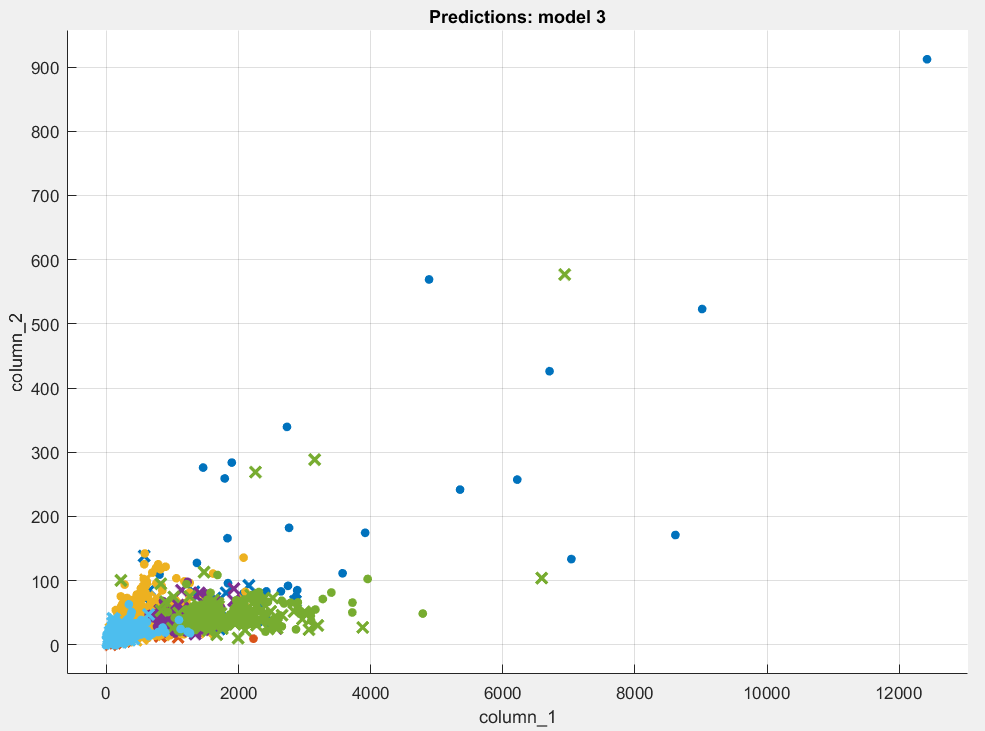


ج) Confusion Matrix



**سابجکت سوم:**

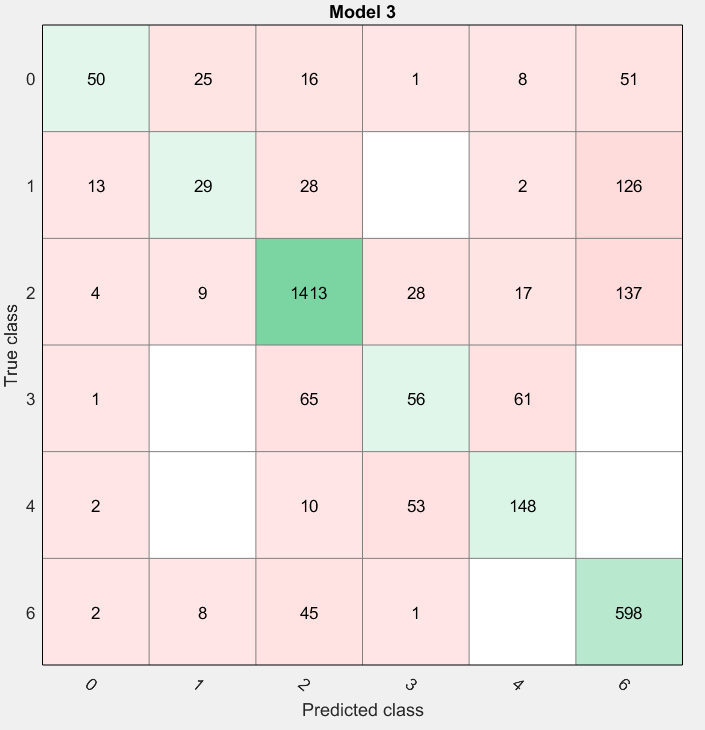
الف) نمودار پراکندگی در دو راستا (ضربدر ها نقاطی هستند که اشتباه تشخیص داده شده اند)



ب) حاصل Cross Validation

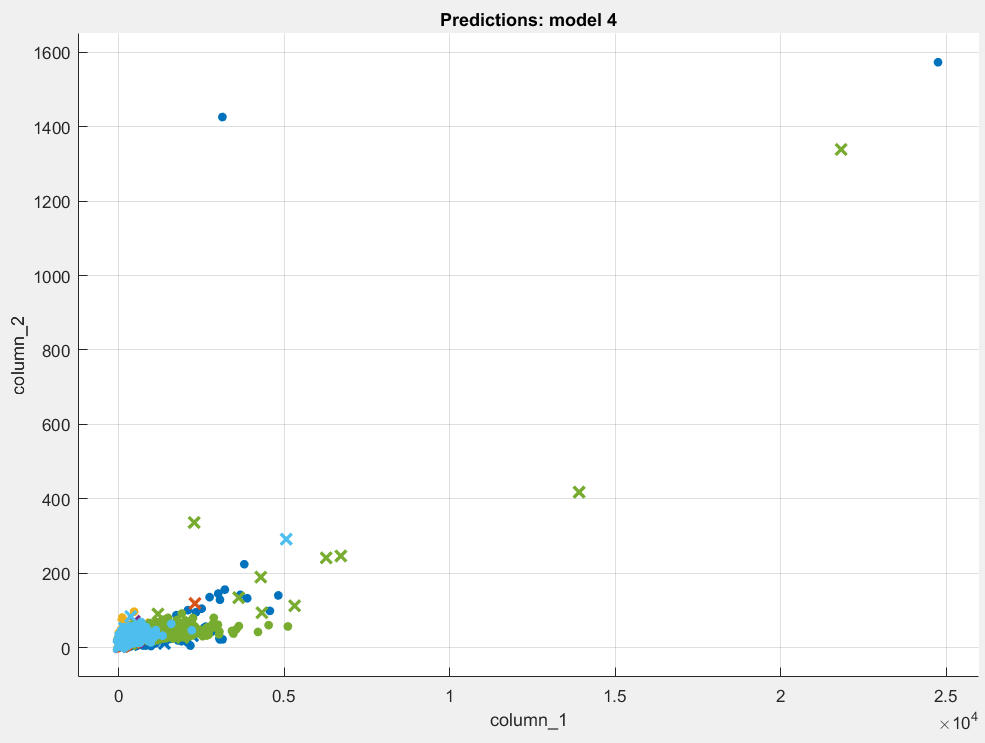


ج) Confusion Matrix



**سابجکت چهارم:**

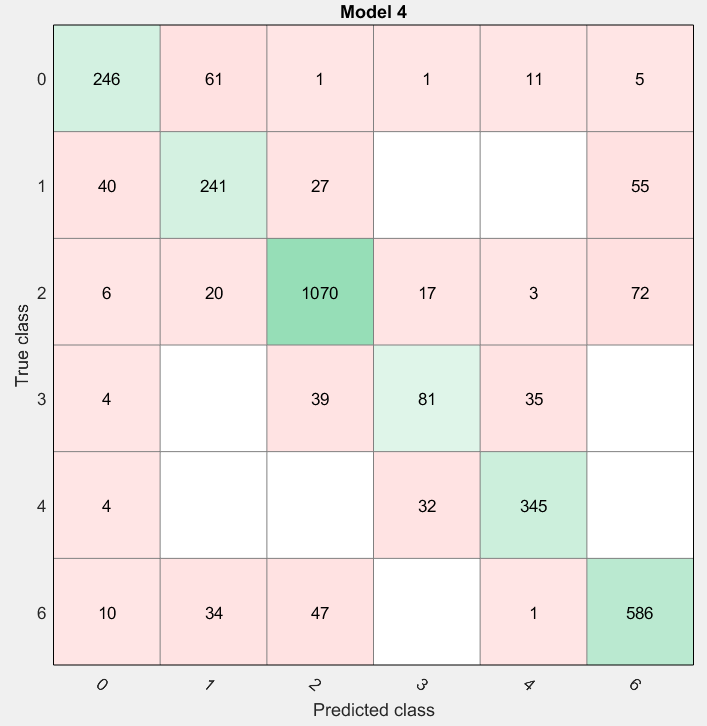
الف) نمودار پراکندگی در دو راستا (ضربدر ها نقاطی هستند که اشتباه تشخیص داده شده اند)



ب) حاصل Cross Validation

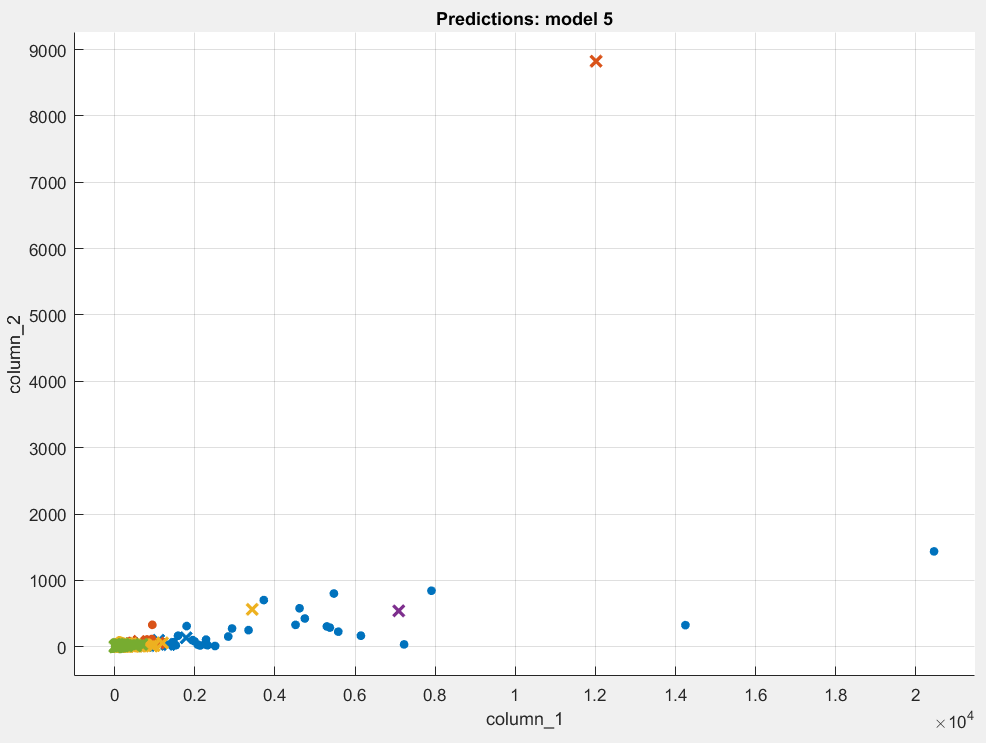


ج) Confusion Matrix



**سابجکت پنجم:**

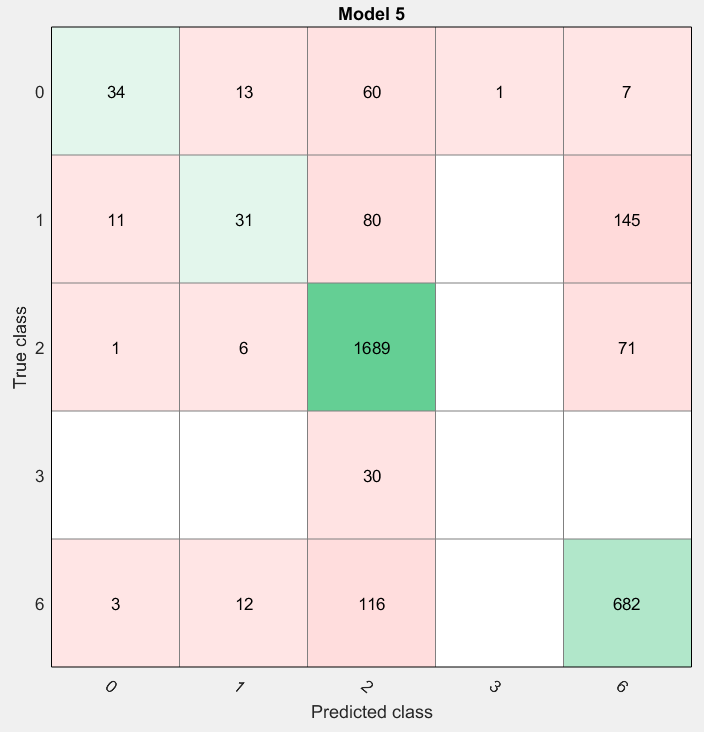
الف) نمودار پراکندگی در دو راستا (ضربدر ها نقاطی هستند که اشتباه تشخیص داده شده اند)



ب) حاصل Cross Validation



ج) Confusion Matrix



**نتیجه‎گیری:**

1- متوسط درصد پاسخ‎گویی (که توسط 5-fold Cross Validation محاسبه شده است) برابر 77%، و محدوده تغییرات درصد پاسخ‌‎گویی بین 72 تا 83 درصد است، که این یعنی درصد پاسخگویی بالا و قابل قبول است، و می‎توان ادعا کرد که 10 ویژگی استخراج‎شده، توانایی جداسازی نسبتاً خوبی برای 6 حالت خواب (و بیداری) دارند.

2- با توجه به Confusion Matrixها، می‎بینیم که داده‎های عمق سوم و چهارم خواب، به شکل معنی‎داری از داده‎های عمق اول و دوم، و نیز حالت REM و بیداری جدا می‎شوند، به این معنی که میزان اشتباه میان این دو دسته (یعنی وضعیتی که داده‎ای متعلق به یکی از این دو دسته باشد و به اشتباه، جزء دسته دیگر شناخته شود) بسیار ناچیز است، اما میزان اشتباه در داخل هر یک از این دو دسته، به نسبت زیاد است.

3- یک نکته‎ی دیگر که جالب توجه است، آن است که این روش دسته‎بندی، عملکرد قابل توجهی برای شناسایی عمق دوم خواب دارد، به طوری که در هر پنج سابجکت، درصد پاسخ صحیح برای این عمق، با اختلاف از سایرین بیشتر است.

## سؤال دوم) خوشه‎بندی

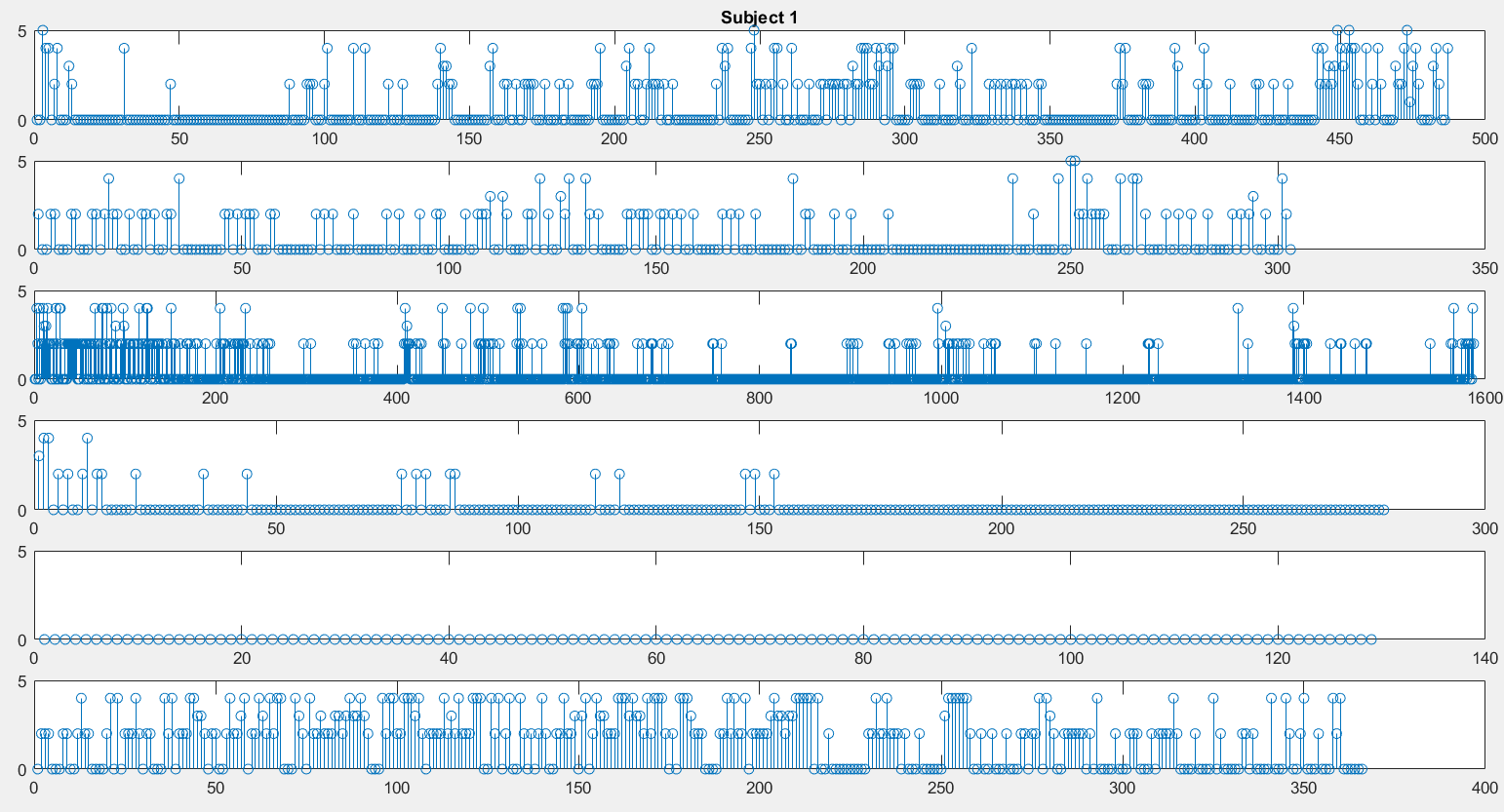
در این بخش می‎خواهیم دیتای به دست آمده از استخراج ویژگی‎های مربوطه را خوشه‎بندی کنیم. این کار که نوعی unsupervised learning است، از جهاتی مشابه بخش قبل است، اما در این جا ما صرفا دیتا را در اختیار الگوریتم قرار می‎دهیم تا آن را به چند دسته تقسیم کند.

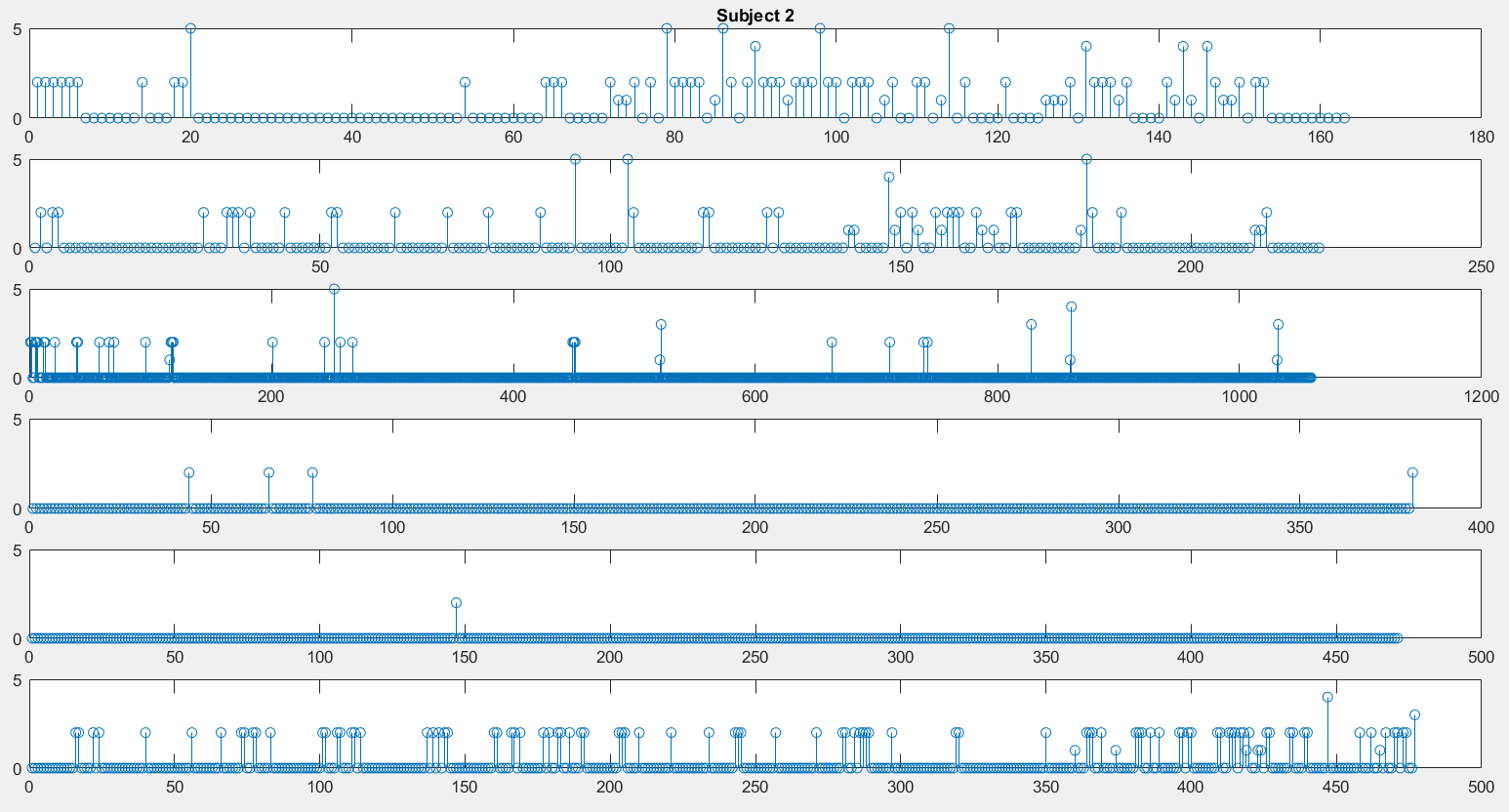
از الگوریتم k-means استفاده می‎کنیم و ماتریس ویژگی‎ها (X) را به عنوان ورودی در نظر می‎گیریم. همچنین با توجه با دانش اولیه خود، k = 6 را انتخاب می‎کنیم.

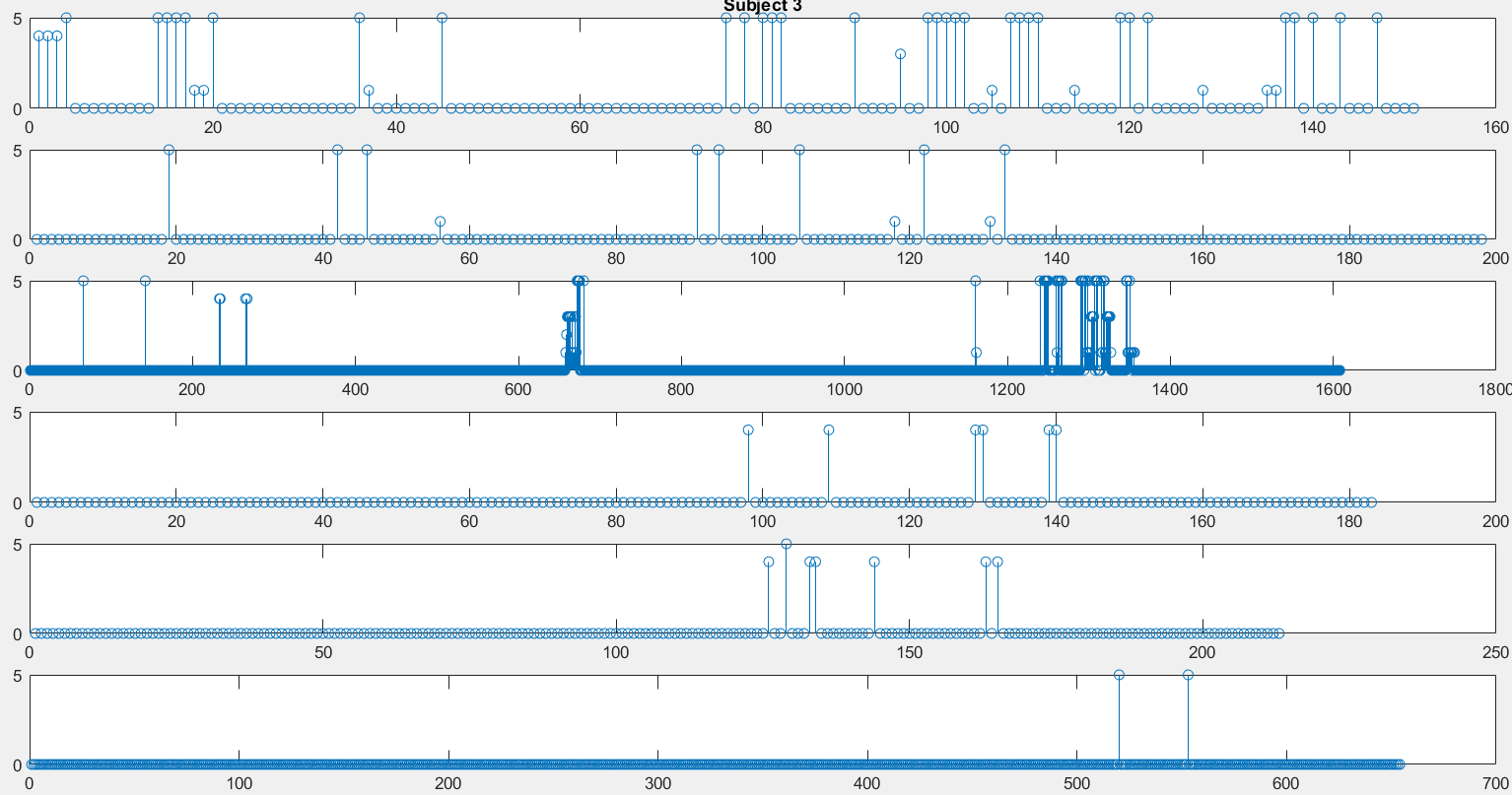
پس از انجام خوشه‎بندی، 6 نمودار رسم می‎کنیم که در هر کدام، شماره گروه جدید (حاصل از خوشه‎بندی) برای اعضای یک گروه (که از قبل می‎دانیم) رسم شده است.

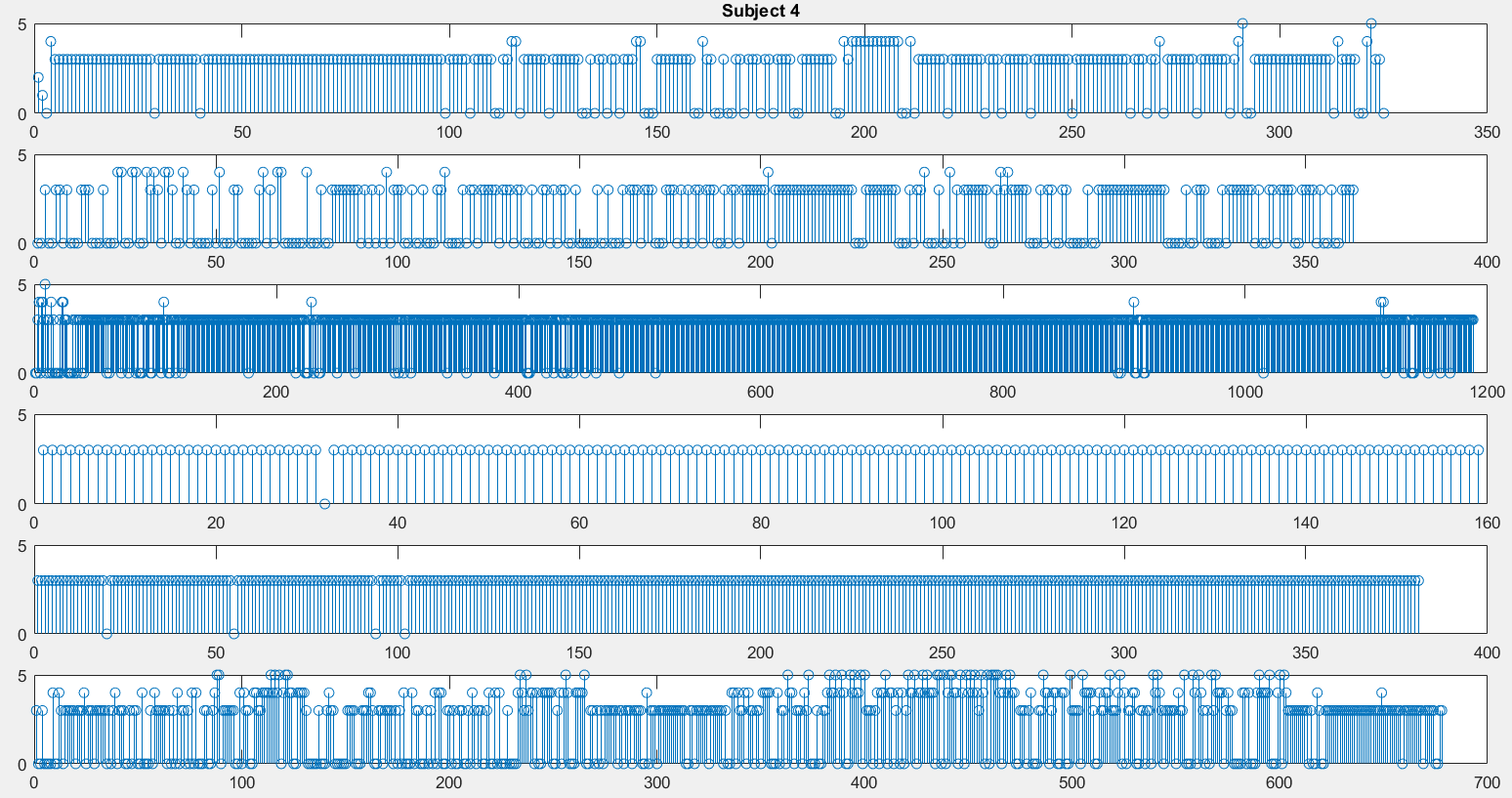
اگر پاسخ کاملاً ایده‎آل باشد، باید نمودارهایی داشته باشیم که هر کدام، به شکل خط افقی y = k هستند که مقدار k در هر نمودار متفاوت است. (این یعنی گروه‎ها دقیقاً طبق انتظارات اولیه ما جدا شده‎اند)

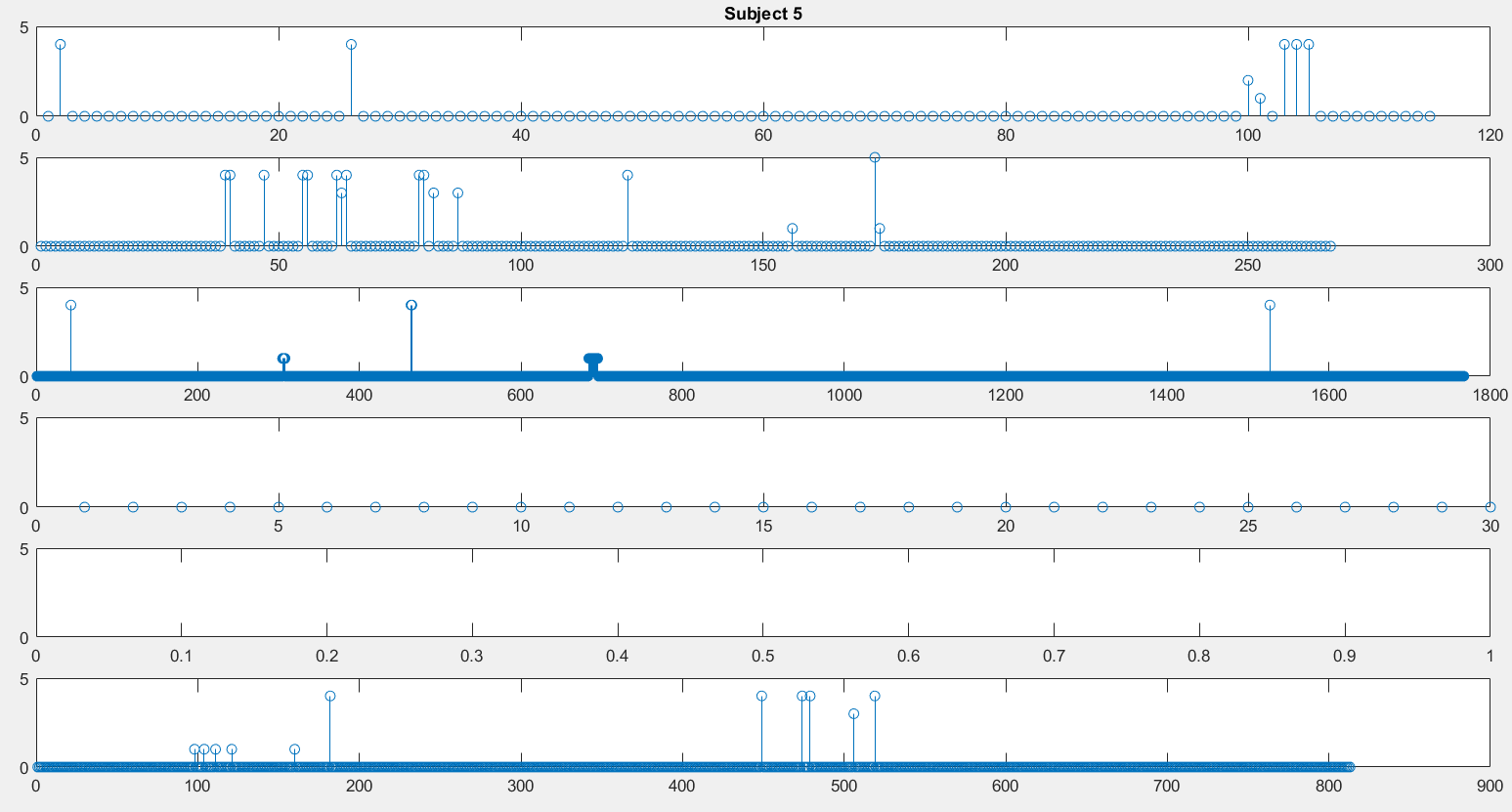
در ادامه نمودارهای متناظر با خوشه‏بندی را برای هر پنج سابجکت مشاهده می‎کنید:











آن‎چه در نمودارهای فوق مشاهده می‎شود، اطلاعات جالبی نیست، چرا که تقریباً در هر شکل، همه‎ی نمودارها نشان‎دهنده‎ی یک مقدار غالب هستند، و سایر مقادیر به شکل پراکنده مشاهده می‎شوند.

به طور خلاصه باید گفت که پاسخ به این سؤال که **«آیا با ویژگی استخراج شده، و با استفاده از روش خوشه‎بندی، می‎توان عمق‎های مختلف خواب و نیز حالت REM و بیداری را از هم جدا کرد؟»** با روشی که ما عمل کردیم، منفی است.

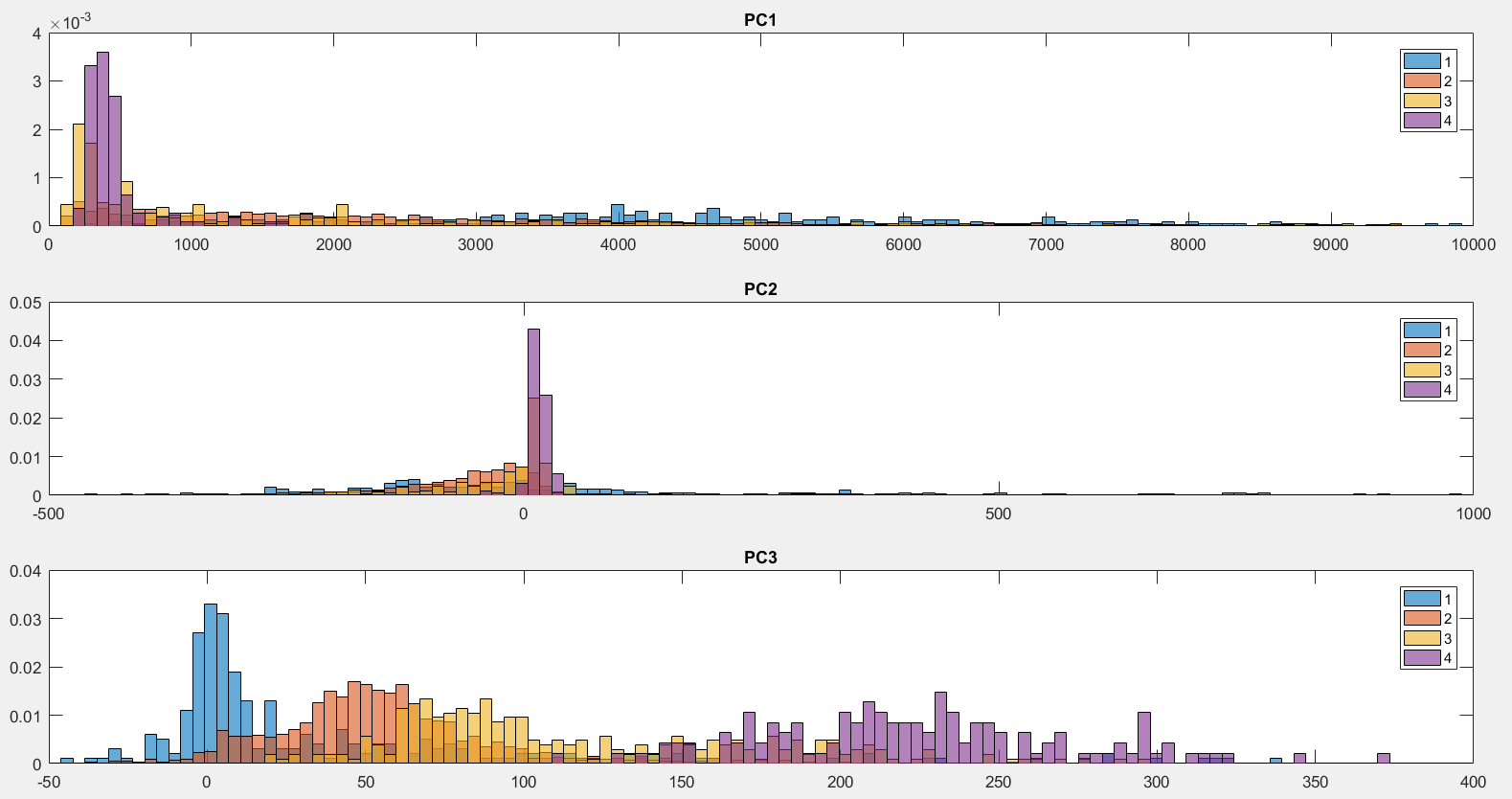
همچنین این ایده را روی ماتریس X که فقط شامل 4 عمق خواب NREM باشد نیز تست کردیم، و نیز سعی کردیم در حالت ساده‎تر، صرفاً دیتای عمق 1 و 2 را از 3 و 4 جدا کنیم (که قبلاً مشاهده کرده بودیم به آسانی انجام می‎شود) که موفقیت‎آمیز نبود.

## سؤال سوم) آیا حافظه‎دار در نظر گرفتن وضعیت‎های خواب اثر مثبتی بر تشخیص عمق خواب دارد؟ به عبارت دیگر، اگر عمق خواب، نه تنها تابعی از ویژگی‎های استخراج شده در همان بازه، بلکه تابعی از زمان‎های قبل نیز باشد، آیا می‎توان به نحو بهتری وضعیت‎های خواب را از هم جدا کرد؟

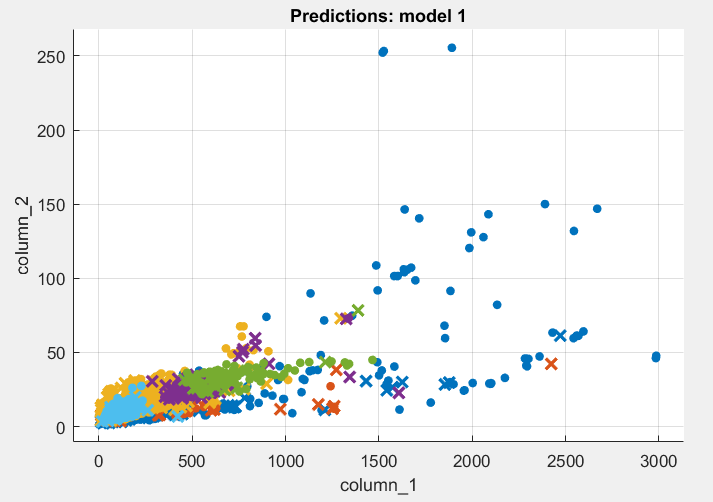
برای پاسخ به این سؤال، ماتریس ویژگی جدیدی را تشکیل می‎دهیم که در آن، برای هر ستون، میانگینی از k بازه 10 ثانیه‎ای قبل گرفته شده و در جای آن ستون قرار داده شده است. در این جا، k را برابر با 5 در نظر می‎گیریم.

با انجام این تغییر در ماتریس X، می‎توانیم تمامی آن‎چه که از ابتدای این گزارش نوشته‎ایم را مجدداً انجام دهیم و نتایج را مقایسه کنیم. اما برای جلوگیری از طولانی شدن (و البته به دلیل کمبود وقت) صرفاً دو بخش رسم هیستوگرام‎های بخش PCA که از روی آن، میزان جداکنندگی را برای عمق‎های چهارگانه‎ی خواب شناسایی می‎کردیم، و بخش SVM (که در سؤال اول قسمت پنجم مطرح کردیم) را مجدداً انجام می‎دهیم.

**سابجکت اول**



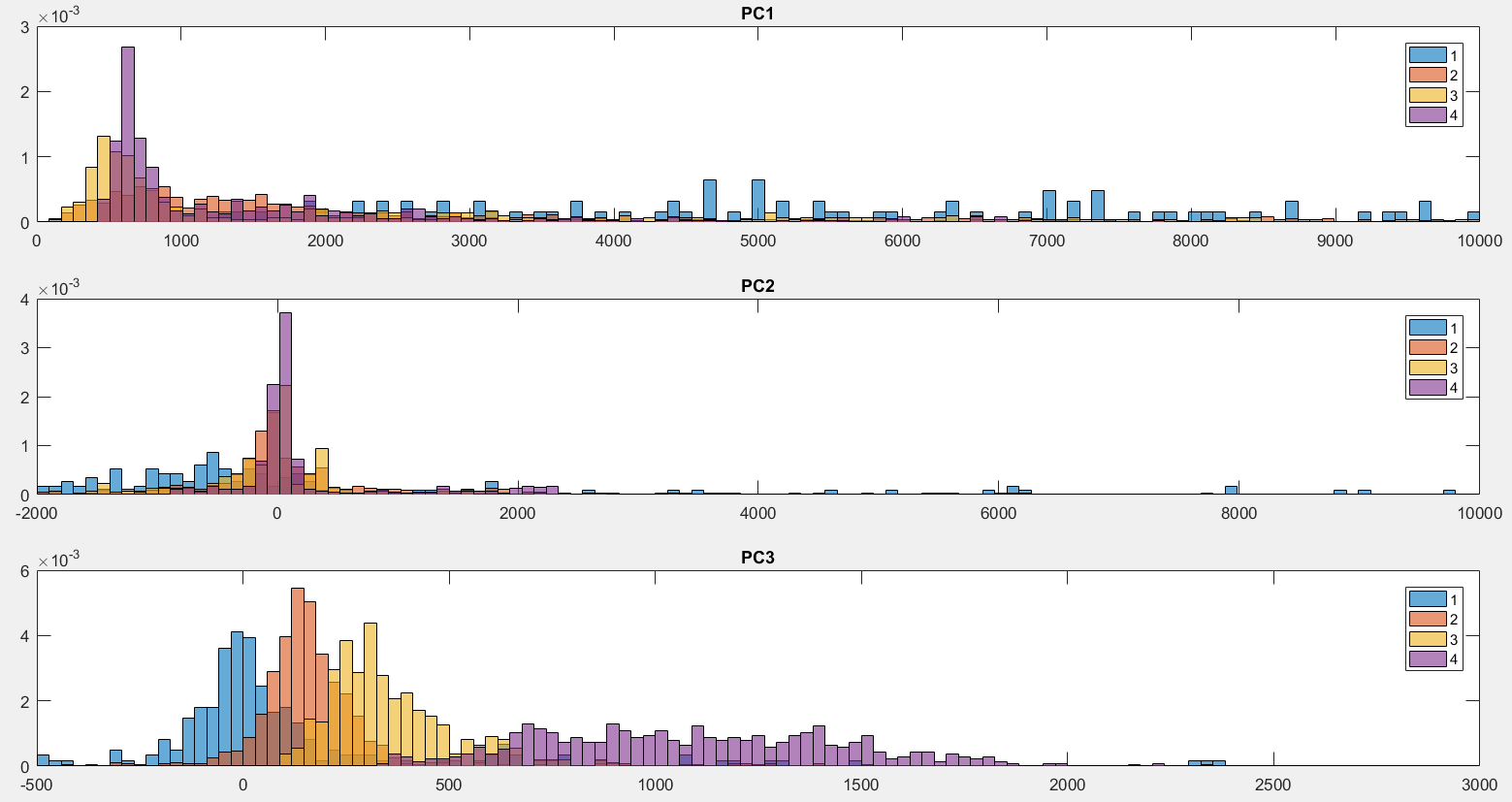
در این نمودار، مشاهده می‎شود که راستای جداکننده (PC3) به شکل قابل توجهی، بهتر عمل می‎کند و لایه‎های خواب را بهتر جدا می‎کند. (برای مقایسه می‎توانید به نمودار متناظر در بخش مربوط به PCA مراجعه کنید)



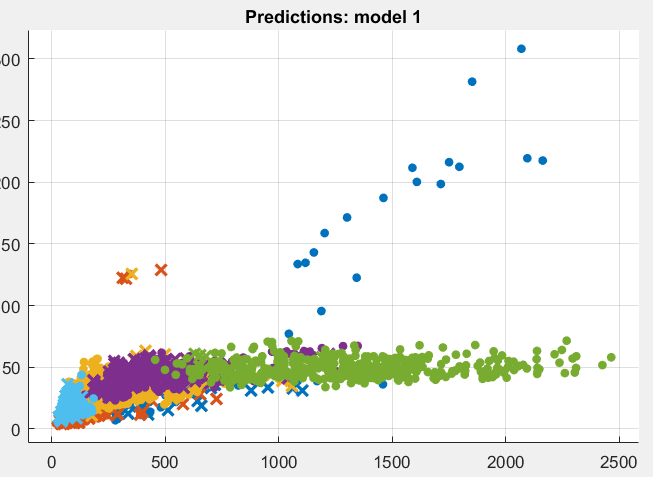


مشاهده می‎شود که درصد جواب صحیح Cross Validation به اندازه 4.8 % درصد افزایش یافته است.

**سابجکت دوم**



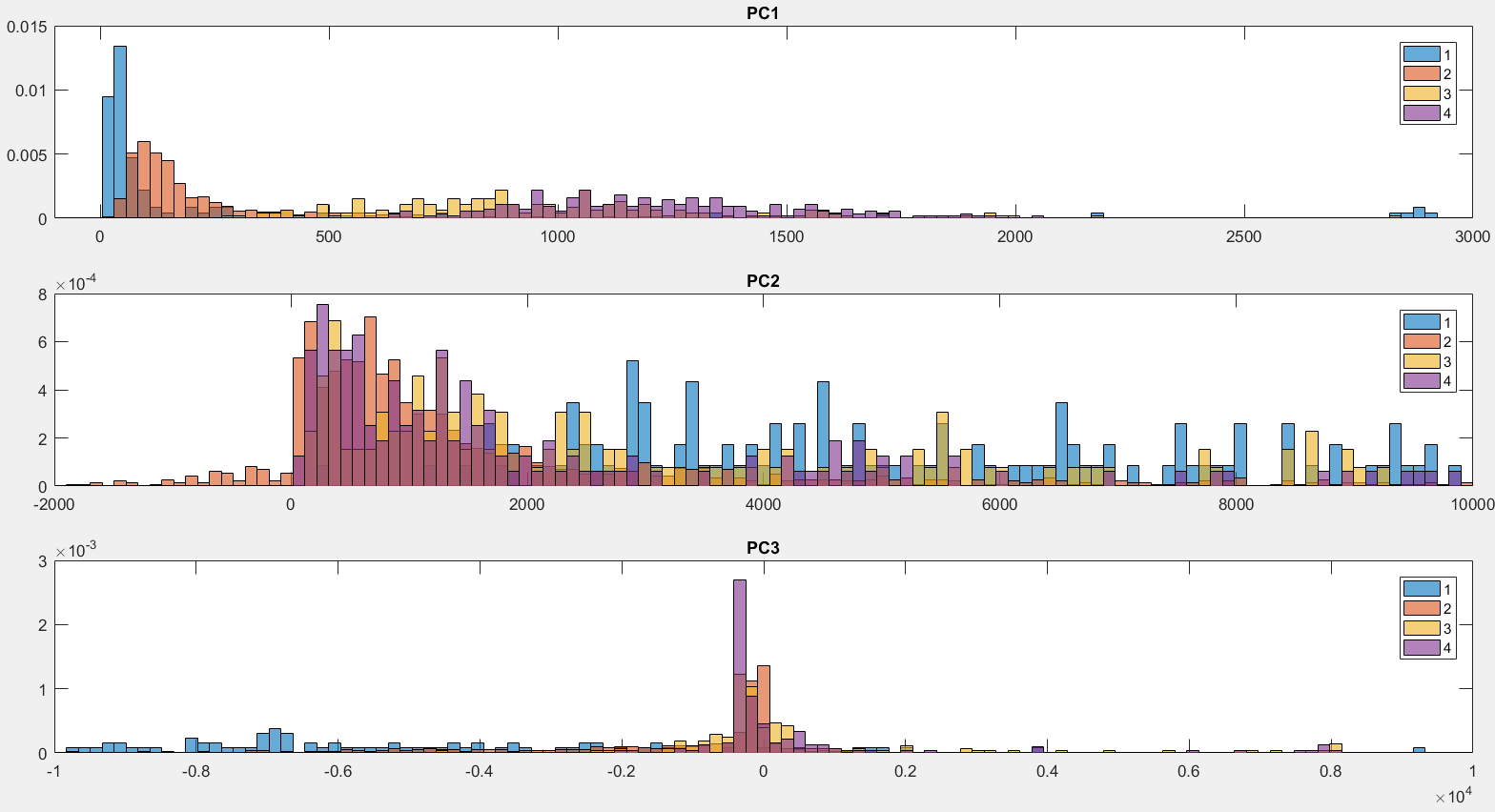
در این نمودار، مشاهده می‎شود که راستای جداکننده (PC3) به شکل قابل توجهی، بهتر عمل می‎کند و لایه‎های خواب را بهتر جدا می‎کند. (برای مقایسه می‎توانید به نمودار متناظر در بخش مربوط به PCA مراجعه کنید)



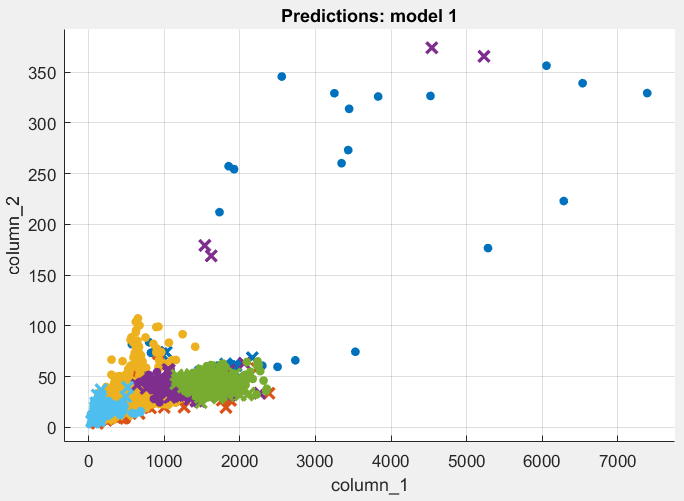


مشاهده می‎شود که درصد جواب صحیح Cross Validation به اندازه 10.5 % درصد افزایش یافته است.

**سابجکت سوم**



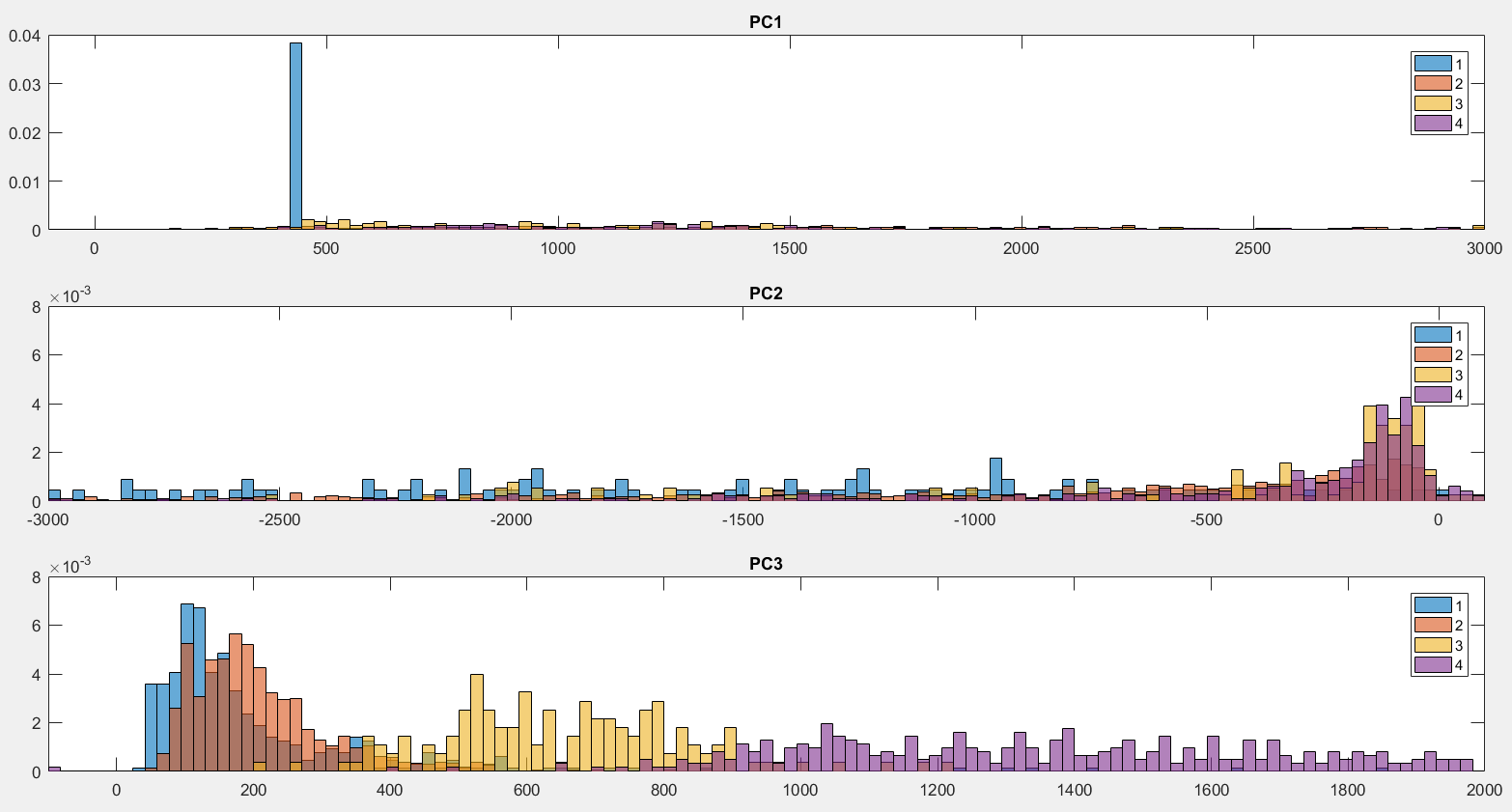
در این نمودار، مشاهده می‎شود که راستای جداکننده (PC1) به شکل قابل توجهی، بهتر عمل می‎کند و لایه‎های خواب را بهتر جدا می‎کند. (برای مقایسه می‎توانید به نمودار متناظر در بخش مربوط به PCA مراجعه کنید)



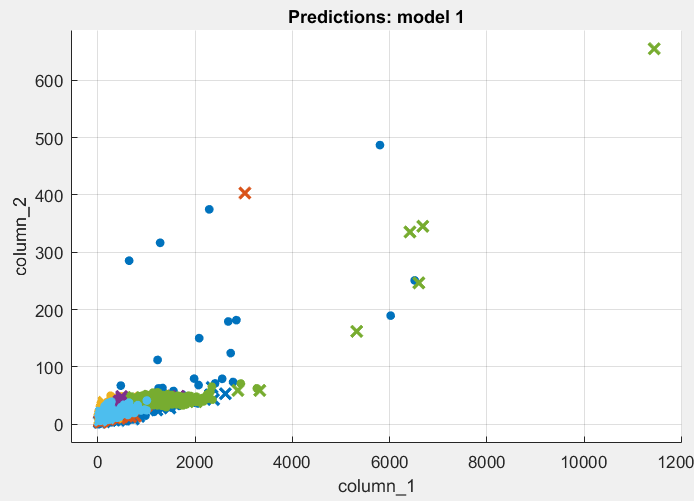


مشاهده می‎شود که درصد جواب صحیح Cross Validation به اندازه 5.8 % درصد افزایش یافته است.

**سابجکت چهارم**



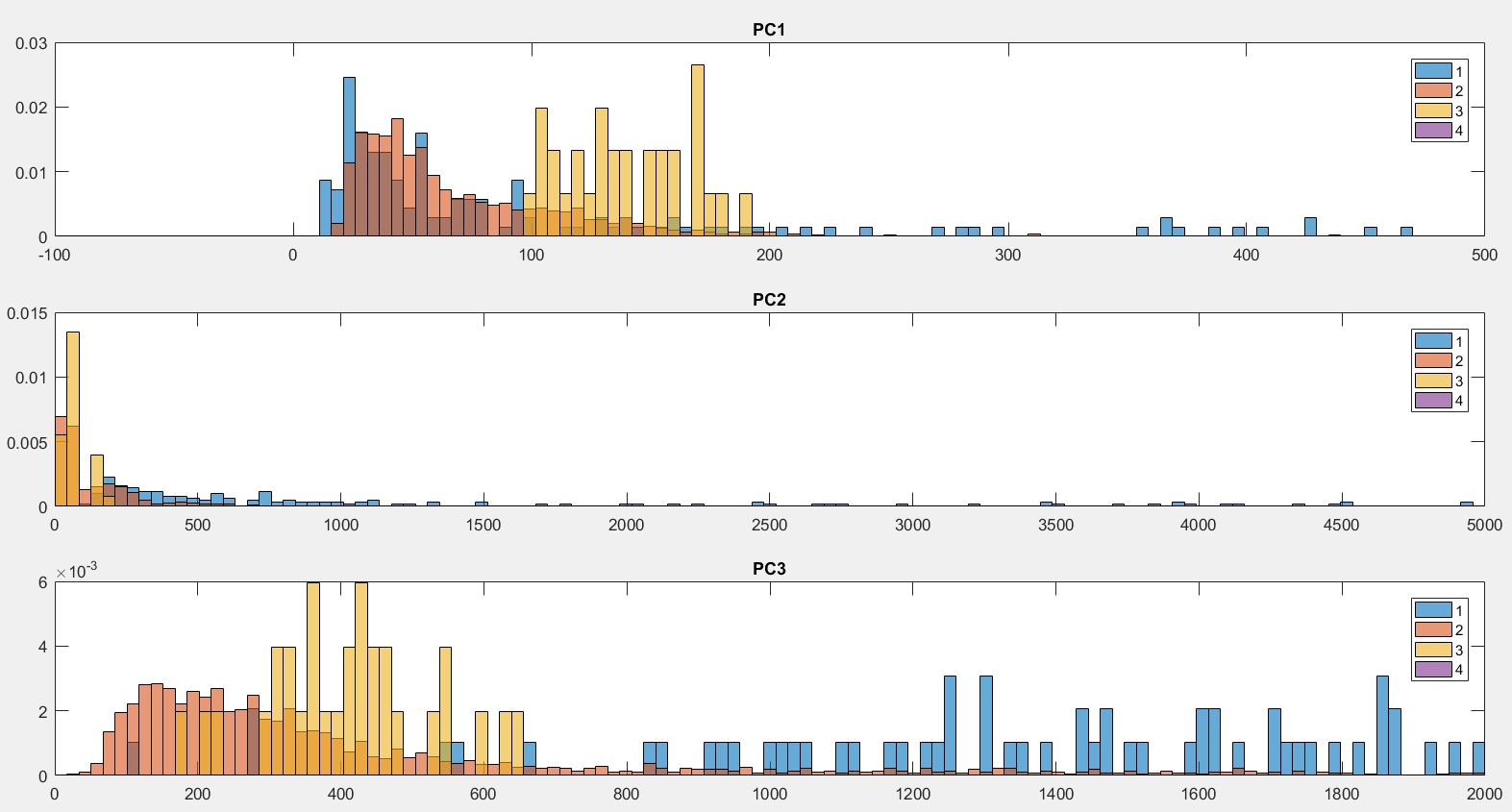
در این نمودار، مشاهده می‎شود که راستای جداکننده (PC3) به شکل قابل توجهی، بهتر عمل می‎کند و لایه‎های خواب را بهتر جدا می‎کند. (برای مقایسه می‎توانید به نمودار متناظر در بخش مربوط به PCA مراجعه کنید)



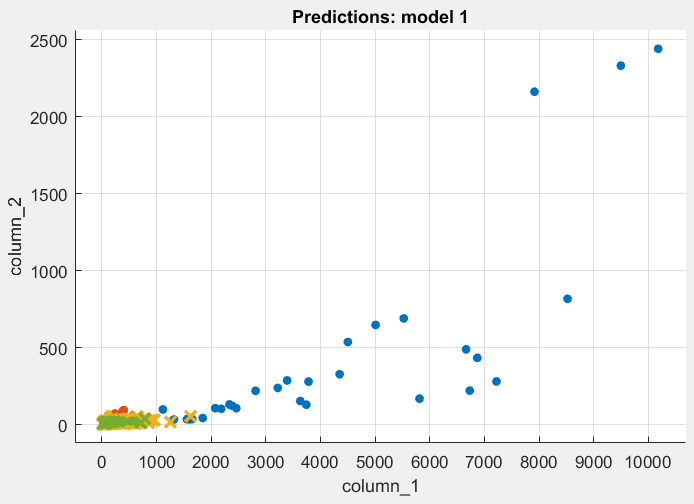


مشاهده می‎شود که درصد جواب صحیح Cross Validation به اندازه 5.6 % درصد افزایش یافته است.

**سابجکت پنجم**



در این نمودار، مشاهده می‎شود که راستای جداکننده (PC1) به شکل قابل توجهی، بهتر عمل می‎کند و لایه‎های خواب را بهتر جدا می‎کند. (برای مقایسه می‎توانید به نمودار متناظر در بخش مربوط به PCA مراجعه کنید)





مشاهده می‎شود که درصد جواب صحیح Cross Validation به اندازه 7.7 % درصد افزایش یافته است.

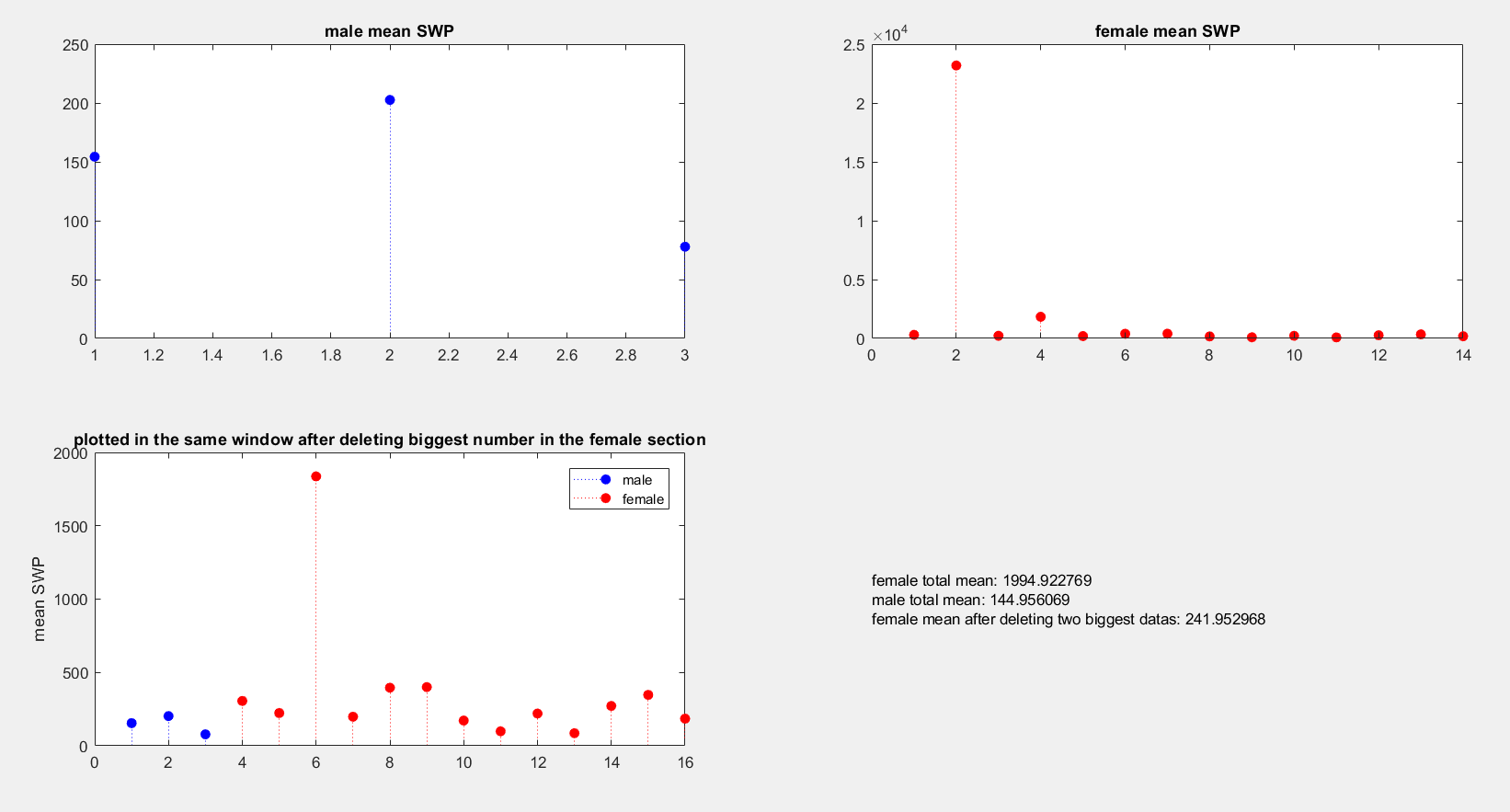
**نتیجه‎گیری:**

همان‎طور که ذکر شد، به دلیل کمبود وقت، صرفاً دو نمونه از فعالیت‎های انجام شده در طول پروژه را با فرض حافظه‎دار بودن وضعیت خواب تکرار کردیم، که در تمامی موارد شاهد افزایش چشمگیری در پاسخگویی (هم در جداسازی لایه‎های خواب توسط PCA، و هم در درصد پاسخ‎گویی Cross Validation) شاهد بودیم که حتی یک نمونه هم پسرفت مشاهده نشد.

بنابراین تقریباً با قطعیت نتیجه می‎گیریم که **وضعیت خواب در هر بازه (10 ثانیه‎ای)، تنها به پارامترهای سیگنال EEG در آن 10 ثانیه وابسته نیست، بلکه عملکردی حافظه‎دار دارد و باید چند بازه زمانی قبل‎تر را نیز در نظر گرفت، تا بتوان توصیف دقیق‎تر و بهتری از عمق خواب ارائه داد.**

## سؤال چهارم) مقایسه تاثیر جنسیت در اندازه SWP

برای بررسی تاثیر جنسیت بر روی SWP ابتدا نیاز به جداسازی این داده ها از الکترود های موجود بود.SWP بنا بر تعریف شامل باند فرکانسی 0.5 تا 2 هرتز می باشد. با اعمال این فیلتر بر روی 2 الکترود موجود برای هر فرد و میانگین گیری، در نهایت عددی را برای مقایسه بدست آوردیم. از آن جا که در داده های اولیه که در گزارش کار ذکر شده بودند تنها 5 فرد موجود بودند و در بین این 5 فرد تنها یک مرد وجود داشت، داده های اضافی را از سایت دانلود کرده ایم. لیست تمامی این افراد در آخر آمده است. البته مشکل تنوع بعد از این افزایش داده نیز بهبود نیافت و در بین 17 نفر، 3 مرد و بقیه زن بودند.



اگر بخواهیم بر همین تعداد داده ی کم تکیه کنیم و نتیجه گیری بکنیم، میتوان گفت که مقدار SWP برای زن ها بیشتر از مرد ها می باشد، چرا که با وجود کم کردن داده های برزگ از سری داده های زن ها، باز هم دارای میانگینی بیشتر از مرد ها میباشند.

2 نکته ی دیگر از نموار قابل استخراج است، یکی اینکه ادعا ی مقاله بر اینکه SWP تغییرات زیادی دارد و از عوامل مختلفی همچون جنسیت و آناتومی سر تاثیر میپذیرد، صحت دارد. اما نکته ی دیگر اینکه انتظار اینکه برای تمامی افراد، SWP زن ها بیشتر از مرد ها باشد، غلط می باشد و به نظر می رسد همواره برقرار نیست. البته در بررسی بالا اثر سن افراد دخیل داده نشده است و برای بررسی دقیق تر ،علاوه بر تنوع داده ی بیشتر، نیاز داریم سن افراد را نیز وارد معادلات بکنیم.

لازم به ذکر است که تمامی این داده ها از افراد در هنگام خواب معمولی و بدون داروی خواب گرفته شده است و نگرانی از اثر داروی خواب بر SWP وارد نیست.

لیست تمامی افراد استفاده شده در این سری داده ها: (پنج نفر اول همان پنج نفر ذکر شده در دستور کار می باشند.)

ST7011J0

ST7022J0

ST7041J0

ST7052J0

ST7061J0

ST7082J0

ST7101J0

ST7112J0

ST7151J0

ST7162J0

ST7171J0

ST7182J0

ST7192J0

ST7201J0

ST7212J0

ST7221J0

ST7241J0