



Caracterización de Fatiga Laboral a Través de Señales Psicofisiológicas

Informe VII

Informe final

Elaborado por: Jorge Gaete
Ángel Jiménez
Cristián Retamal
Ignacio Vargas
Marcelo Riquelme
Sebastián Tello
Fecha: 31 de Marzo, 2019

Tabla de Contenidos

INTRODUCCIÓN	7
DESCRIPCIÓN DEL PROYECTO	8
OBJETIVOS DEL INFORME	9
1.- REVISIÓN BIBLIOGRÁFICA	10
FATIGA MENTAL Y FÍSICA	10
ENFOQUES DE MEDICIÓN DE FATIGA FÍSICA Y MENTAL	11
SEÑALES PSICOFISIOLÓGICAS	12
ANÁLISIS ESTADÍSTICO Y APRENDIZAJE DE MÁQUINA	14
CONCLUSIONES DEL ESTADO DEL ARTE	15
2.- METODOLOGÍA PARA LA PREDICCIÓN DE FATIGA	17
A.- DEFINICIÓN DE LA TAREA	18
1.- CARACTERIZACIÓN DE LA TAREA	18
2.- CARACTERIZACIÓN DEL ENTORNO DE EJECUCIÓN	19
3.- DEFINICIÓN DE LA MUESTRA	20
4.- PREPARACIÓN DE ARTEFACTOS	20
B.- DEFINICIÓN DEL EXPERIMENTO	23
5.- DEFINICIÓN DEL PROTOCOLO EXPERIMENTAL	23
6.- EJECUCIÓN DEL PROTOCOLO EXPERIMENTAL	24
C.- PREPARACIÓN DE DATOS	24
7.- LIMPIEZA Y PROCESAMIENTO DE DATOS	25
8.- ETIQUETADO DE FATIGA	30
D.- ANÁLISIS DE LA TAREA	34
9.- EFECTO DE LA TAREA POR INDIVIDUO	34
10.- DESCRIPCIÓN DE LA TAREA PARA LA MUESTRA AGREGADA	36
11.- RESULTADOS Y CONCLUSIONES	38
3.- APLICACIÓN DE LA METODOLOGÍA: CONDUCCIÓN	39
A.- DEFINICIÓN DE LA TAREA	39
1.- CARACTERIZACIÓN DE LA TAREA	39
2.- CARACTERIZACIÓN DEL ENTORNO DE EJECUCIÓN DE LA TAREA	41
3.- DEFINICIÓN DE LA MUESTRA	47
4.- PREPARACIÓN DE ARTEFACTOS Y MATERIALES	48
B.- DEFINICIÓN DEL EXPERIMENTO	48
5.- DEFINICIÓN DEL PROTOCOLO EXPERIMENTAL	48
6.- EJECUCIÓN DEL PROTOCOLO EXPERIMENTAL	49
C.- PREPARACIÓN DE DATOS	51
7.- LIMPIEZA Y PROCESAMIENTO DE DATOS	51

8.- ETIQUETADO DE FATIGA	52
D. ANÁLISIS DE LA TAREA	60
9.- EFECTO DE LA TAREA POR INDIVIDUO	60
10.- DESCRIPCIÓN DE LA TAREA PARA LA MUESTRA AGREGADA	62
11.- RESULTADOS Y CONCLUSIONES	70
4.- ANÁLISIS COMPARATIVO DE LA APLICACIÓN DE LA METODOLOGÍA	72
A.- DEFINICIÓN DE LA TAREA	73
1.- CARACTERIZACIÓN DE LA TAREA	73
2.- CARACTERIZACIÓN DEL ENTORNO DE EJECUCIÓN	74
3.- DEFINICIÓN DE LA MUESTRA	75
4.- PREPARACIÓN DE ARTEFACTOS Y MATERIALES	75
B.- DEFINICIÓN DEL EXPERIMENTO	76
5.- DEFINICIÓN DEL PROTOCOLO EXPERIMENTAL	76
6.- EJECUCIÓN DEL PROTOCOLO EXPERIMENTAL	76
C.- PREPARACIÓN DE DATOS	78
7.- LIMPIEZA Y PROCESAMIENTO DE DATOS	78
8.- ETIQUETADO DE FATIGA	79
D. ANÁLISIS DE LA TAREA	80
9.- EFECTO DE LA TAREA POR INDIVIDUO	80
10.- DESCRIPCIÓN DE LA TAREA PARA LA MUESTRA AGREGADA	80
11.- RESULTADOS Y CONCLUSIONES	81
5.- ANÁLISIS EXPLORATORIO DE LA PREDICCIÓN DE FATIGA USANDO MÉTODOS DE DEEP LEARNING	82
CONCLUSIONES Y PRÓXIMOS PASOS	86
CONCLUSIONES DEL INFORME	86
PRÓXIMOS PASOS	87
ANEXO I: GRÁFICO DE SELECCIÓN DE CARACTERÍSTICAS CON ALGORITMO BORUTA PARA CADA SUBACTIVIDAD	93
ANEXO II: TABLA DE LA VARIANZA EXPLICADA POR LAS COMPONENTES PRINCIPALES PARA CADA SUBACTIVIDAD	95
ANEXO III: VALORES BIC PARA DECIDIR KERNEL Y NUMERO DE GRUPOS DE SUJETOS, PARA CADA SUBACTIVIDAD	96
ANEXO IV: VALORES DEL PUNTAJE DE SILHOUETTE PARA DETERMINAR LA CANTIDAD DE GRUPOS DE SUJETOS, PARA CADA SUBACTIVIDAD	98
ANEXO V: VALORES BIC DE LAS ITERACIONES DEL MODELO DE AGRUPACIÓN GMM PARA CADA SUBACTIVIDAD	100
ANEXO VI: PRUEBA DE LEVENE PARA LA VALIDACIÓN DE AGRUPACIÓN DE SUJETOS EN CONDUCCIÓN, IDA CIUDAD	101
ANEXO VII: PRUEBA DE LEVENE PARA LA VALIDACIÓN DE AGRUPACIÓN DE SUJETOS EN CONDUCCIÓN, IDA MONTAÑA	102
ANEXO VIII: PRUEBA DE LEVENE PARA LA VALIDACIÓN DE AGRUPACIÓN DE SUJETOS EN CONDUCCIÓN, VUELTA MONTAÑA	103
ANEXO IX: PRUEBA DE LEVENE PARA LA VALIDACIÓN DE AGRUPACIÓN DE SUJETOS EN CONDUCCIÓN, VUELTA CIUDAD	104

ANEXO X: PRUEBA F PARA LA VALIDACIÓN DE AGRUPACIÓN DE SUJETOS EN CONDUCCIÓN, IDA CIUDAD	105
ANEXO XI: PRUEBA F PARA LA VALIDACIÓN DE AGRUPACIÓN DE SUJETOS EN CONDUCCIÓN, IDA MONTAÑA	106
ANEXO XII: PRUEBA F PARA LA VALIDACIÓN DE AGRUPACIÓN DE SUJETOS EN CONDUCCIÓN, VUELTA CIUDAD	107
ANEXO XIII: PRUEBA F PARA LA VALIDACIÓN DE AGRUPACIÓN DE SUJETOS EN CONDUCCIÓN, VUELTA MONTAÑA	108
ANEXO XIV: INTERPRETACION SEGÚN NIVELES DE SIGNIFICANCIA	109
ANEXO XV: TABLA DE CONTINGENCIA DE FATIGA SNA VS GRUPOS PARA CADA SUBACTIVIDAD	110
ANEXO XVI: TABLA DE CONTINGENCIA DE FATIGA DE RENDIMIENTO VS GRUPOS PARA CADA SUBACTIVIAD	112
ANEXO XVII: PRUEBA DE LEVENE, FATIGA SNA	113
ANEXO XVIII: PRUEBA DE LEVENE, FATIGA RENDIMIENTO	114
ANEXO XIX: EJEMPLO DE INTERPRETACIÓN DE LA TENDENCIA PARA LA FATIGA SNA Y DE RENDIMIENTO	115

RESUMEN EJECUTIVO

En este informe final se presenta la validación de la metodología para el proceso de captura, procesamiento y análisis de datos psicofisiológicos para la caracterización de la fatiga producida por tareas laborales en trabajadores. Esto en el contexto del proyecto “caracterización de fatiga física y mental asociada a tareas laborales a través del sensado y procesamiento de señales psicofisiológicas”, desarrollado conjuntamente por el IST y la Universidad de Chile. Específicamente este informe tiene como objetivo el describir la validación de la metodología en una tarea laboral real, incluyendo un análisis comparativo de los resultados obtenidos de la aplicación en ambientes controlados y en terreno. En este documento se presentan los siguientes puntos:

1. Introducción del proyecto.
2. Presentación del trabajo relacionado.
3. Descripción de la metodología de captura, procesamiento y análisis de datos mejorada.
4. Aplicación de la metodología en una actividad laboral en terreno.
5. Validación de la metodología mediante un análisis comparativo.
6. Análisis exploratorio de la predicción de fatiga.
7. Conclusiones sobre la aplicación de la metodología y presentación del trabajo futuro.

En las entregas anteriores del proyecto se ha avanzado en la construcción de una metodología para la caracterización de las tareas laborales en función del estrés que éstas generan sobre los trabajadores y su aplicación en entornos controlados. Esta metodología surge desde un acabado análisis de la literatura, del trabajo en campo y de la experimentación con sensores. En esta entrega se destacan los resultados finales de la aplicación de esta metodología incluyendo la validación con experimentos realizados en terreno y el análisis exploratorio de la predicción de fatiga con métodos de deep learning.

INTRODUCCIÓN

La fatiga laboral es cada vez más reconocida mundialmente como un problema de la industria moderna al ser una de las causas principales de los accidentes en el trabajo (Sadeghniaat-Haghighi & Yazdi, 2015; Kant et al, 2003). La Organización Internacional del Trabajo (OIT) estima que las pérdidas económicas por enfermedades y accidentes relacionados con el trabajo son significativas: cercanas a un 4% del PIB mundial, lo que equivale a US\$2,8 billones (Dirección de Estudios Previsionales, 2015).

Según estudios de la Mutual de Seguridad, se comprobó que en la industria latinoamericana los incidentes con daños serios a la integridad de los trabajadores no superan el 2% o 3%. Sin embargo, el resto de los incidentes ocasiona pérdidas, en promedio, del 18% al 20% del presupuesto de las operaciones por efecto de daños a la propiedad, fallas operacionales y desviación del proceso. Además, genera costos económicos por ineficiencias, días de trabajo perdidos, pagos compensatorios, atención médica y rehabilitación, lo que finalmente impacta en pérdidas que reducen la ganancia neta (Sigweb, 2011).

Haciendo frente a esta realidad, el IST y la Universidad de Chile buscan abordar este problema en la industria de servicios mediante la construcción de una metodología para la medición y clasificación de la fatiga sufrida por los trabajadores durante la ejecución de su trabajo. Con esto se busca entregar herramientas metodológicas para el monitoreo y pronóstico de fatiga que apoyen en distintos ámbitos la gestión de la fuerza laboral, tales como en la programación y reconfiguración de turnos o en la prevención de enfermedades psicolaborales y accidentes, entre otros. Esta metodología fue desarrollada y testeada de manera iterativa primero en experiencias en ambientes controlados y finalmente con una tarea laboral real en la industria del transporte de pasajeros. La naturaleza iterativa en la creación de la metodología permite entregar una metodología más general y aplicable a distintas tareas laborales.

Este informe final entrega los resultados de la implementación de la metodología desarrollada en dos experimentos controlados (digitación y reposición) y uno en terreno (conducción). Como principal resultado se tiene que la metodología desarrollada permite la identificación de fatiga en los trabajadores tanto en los entornos controlados como en el real. Esta identificación es no trivial y se logra a partir del procesamiento de señales psicofisiológicas de los sujetos, capturadas de manera no invasiva por medio de biosensores mientras realizan sus tareas.

DESCRIPCIÓN DEL PROYECTO

El presente informe se enmarca dentro de una iniciativa que busca la definición de una metodología para la caracterización de la fatiga experimentada por trabajadores durante el ejercicio de sus tareas laborales. Este proyecto se compone de tres etapas, en las que se realizan una serie de actividades para la construcción iterativa de una metodología replicable para la captura y procesamiento de señales psicofisiológicas, la obtención de patrones de fatiga física y mental experimentada por los trabajadores y una propuesta metodológica final para la caracterización de patrones de fatiga en otras industrias. La tabla I presenta y descompone estas etapas en sus actividades principales.

Etapa y descripción	Actividades principales
A) Modelos para la identificación de la fatiga física y mental	i) Revisión del Estado del arte. ii) Desarrollo de módulos para la captura y almacenamiento de señales psicofisiológicas. iii) Modelos para la detección de fatiga física y mental a partir de señales psicofisiológicas en tiempo real.
B) Caracterización de patrones de fatiga física y mental por actividad laboral	i) Identificación de patrones de fatiga física y mental en actividades laborales en tiempo real. ii) Integración y documentación del proceso de captura, procesamiento y análisis de datos.
C) Formulación y validación de la metodología	i) Formulación de una metodología a partir del proceso integrado de captura, procesamiento y análisis. ii) Aplicación de la metodología y validación con un nuevo caso de uso.

Tabla I: Etapas y actividades principales del proyecto.

Con las etapas A, B y C reportadas en los informes anteriores (informes I, II, III, IV, V y VI), el presente constituye la entrega final de la actividad C) Formulación y validación de la metodología. Específicamente de la actividad C.ii. En este informe se presentan los resultados finales de la aplicación de la metodología presentada en el informe VI, aplicada a tareas laborales simuladas en laboratorio y a un experimento en terreno con los conductores de una empresa privada de transporte de pasajeros al manejar desde Santiago hacia la mina Los Bronces.

OBJETIVOS DEL INFORME

El presente informe entrega el resultado de la aplicación de la metodología para la caracterización fisiológica del fenómeno de fatiga en actividades laborales. En particular, se presenta la metodología final, su aplicación en tres actividades laborales y se contextualiza la metodología en el estado del arte de la investigación.

En particular en este informe se presentan los pasos metodológicos en una actividad laboral realizada en terreno: conductores de buses de una empresa privada de transporte de pasajeros en la ruta Santiago-Los Bronces. Además, se presentan los resultados de la aplicación de las herramientas de análisis de datos para caracterizar la fatiga inducida por esta tarea en los sujetos experimentales. Luego, se realiza un análisis comparativo con dos tareas ejecutadas en laboratorio, las que se detallan en documentos anexos pues fueron tratadas en informes anteriores. Específicamente, se da cuenta de los siguientes objetivos específicos:

- *Presentar la metodología desarrollada y enmarcarla dentro del contexto del desarrollo de metodologías para la caracterización de la fatiga en el entorno laboral.*
- *Presentar los resultados de la aplicación de la metodología para la caracterización de una tarea en función de la fatiga que produce durante su ejecución aplicada en una actividad laboral real y en terreno (conducción).*
- *Validar la metodología aplicada comparando los resultados obtenidos en dos actividades laborales en ambiente controlado (laboratorio: digitación y reposición) versus una actividad laboral real en terreno (conducción).*
- *Concluir de manera final a las hipótesis principales de este trabajo y discutir futuras extensiones del trabajo realizado.*

1.- REVISIÓN BIBLIOGRÁFICA

Fatiga mental y física

La fatiga se entiende en general como cualquier pérdida de eficiencia o como la aversión a cualquier tipo de esfuerzo (Grandjean, 1970). En particular, la fatiga física y mental se definen como el resultado psicofisiológico de periodos prolongados de actividad física y cognitiva que producen una disminución en el desempeño de una tarea (Boksem, 2008). Desde el punto de vista operacional, la fatiga es un factor de reducción de la productividad de una tarea y de la mala calidad de sus resultados (Aryal., et al., 2017).

La fatiga mental en el ámbito laboral se desarrolla cuando los trabajadores se ven expuestos a estrés, el cual se define como cualquier respuesta no específica del cuerpo ante cualquier demanda (Selye, 1956), y que desencadena una serie de reacciones de patrones complejos con componentes fisiológicos, cognitivos y de comportamiento (Feuerstein, et al., 1986). El estrés en el trabajo surge cuando no están bien ajustadas la demanda de tareas con las capacidades de los trabajadores (Van Daalen, et al., 2009). Sin embargo, se ha demostrado que trabajando con baja demanda de recursos cognitivos, se genera aburrimiento y así una tendencia a cometer errores (Ryu and Myung, 2005). Cuando un trabajador se ve expuesto a estrés de forma mantenida en el tiempo, puede desarrollar estrés crónico y depresión (Cinaz et al 2013), (Kim, Y. K., & Cha, N. H., 2015). A su vez, el estrés crónico trae consigo riesgos pato-fisiológicos, entre ellos, enfermedades cardiovasculares, enfermedades cerebrovasculares, diabetes y deficiencias en el sistema inmune (Cacioppo J.T., et al., 2000). Dentro del ambiente laboral y en específico en la atención de clientes, el factor psicosocial más frecuente como fuente de estrés es "lidiar con clientes difíciles" (Lin, Y. H., et al., 2010), (J Wegge, et al., 2007).

La fatiga física, dentro del contexto las actividades laborales de oficina, está asociada a incomodidades en diferentes partes del cuerpo que se generan durante o después de una jornada prolongada de trabajo. Trabajar en un call center durante más de cuatro años está fuertemente asociado con la incomodidad en todas las áreas del cuerpo y otros problemas de salud (Lin, Y. H., et al., 2009). En particular, incomodidad musculoesquelética, fatiga ocular y dolor de garganta son las molestias físicas significativas en los operadores de call centers según Lin, Y.H. et al. (Lin, Y. H., et al., 2010). Este último estudio también concluye que el 33.5% de los operadores de llamadas de salida y el 27.1% de los operadores de llamadas de entrada, experimentan un alto nivel estrés en el trabajo.

Enfoques de medición de fatiga física y mental

Para medir el nivel de fatiga que experimenta un individuo existen tanto enfoques subjetivos como cuantitativos. El primero se basa en la utilización de la carga de trabajo mental, la cual se define como el costo mental o nivel de esfuerzo para lograr los requerimientos de una tarea (Hart & Wickens, 1990). Para medirla se utilizan cuestionarios autoinformados, siendo el NASA-TLX el más difundido dentro de la literatura (Hart & Staveland 1988). Este cuestionario evalúa las siguientes seis subescalas: (1) demanda mental, (2) demanda física, (3) demanda temporal, (4) desempeño, (5) esfuerzo y (6) frustración. Para medir otras dimensiones de la fatiga física y mental se utilizan diferentes cuestionarios tales como: 12-item General Health Questionnaire (GHQ-12; Goldberg & Williams 1988), ampliamente usado para medir estrés; Work-Family Conflict Scale (Carlson et al. 2000) que mide conflictos de demandas incompatibles entre el trabajo de una persona y su vida personal; Positive and Negative Affectivity Schedule (PANAS; Watson et al. 1988), el cual mide las emociones que experimenta un sujeto. Además, se suman los cuestionarios de incomodidades físicas que se utilizaron en los estudios de J. Wegge y en el de Y. Lin.

En el contexto del presente trabajo, donde la idea es medir la reacción fisiológica y de rendimiento de los trabajadores durante la ejecución de las tareas laborales, resulta engorroso aplicar este tipo de cuestionarios durante la ejecución de las actividades. En la literatura se ha optado por medir el estrés con una escala de Likert de uno a cinco, desde no estresante a extremadamente estresante, disminuyendo considerablemente la intrusividad del experimento (J. Hernández, et al., 2011) (J. Allan, et al., 2009), mientras que en el trabajo de Wilson & Russell (2003) para medir la fatiga en operadores de torres de control aéreo se optó por pre-etiquetar el nivel de carga mental que generan diferentes tipos de tareas/llamadas.

El enfoque cuantitativo se basa en el uso de biosensores para registrar señales psicofisiológicas, tales como la frecuencia cardíaca, la conductividad eléctrica de la piel, la dilatación pupilar, el nivel de oxígeno en la sangre, la secreción de adrenalina, entre muchas otras. Por medio del procesamiento analítico de estas señales es posible estimar el efecto psicofisiológico de la demanda física y mental que genera una tarea sobre un trabajador de manera poco invasiva y de bajo costo. Usando estas señales y el nivel de fatiga autoinformado de los sujetos de prueba, es posible utilizar métodos cuantitativos para clasificar estos niveles de fatiga de manera automática y en tiempo real. Este enfoque permite también el uso de métodos estadísticos para estimar el nivel de fatiga que tendrá un trabajador en el mediano plazo (Cavanaugh, et al., 2000) a partir del sensado de señales fisiológicas durante un periodo largo de tiempo.

Señales psicofisiológicas

Las señales psicofisiológicas son manifestaciones periféricas de procesos gobernados por el sistema nervioso. Éste se divide, como se observa en la Figura 1, en el Sistema Nervioso Central (SNC), formado por el encéfalo y la médula espinal, el Sistema Nervioso Somático, encargado de los receptores sensoriales y movimientos voluntarios y el Sistema Nervioso Autónomo (SNA) que regula las acciones involuntarias. En particular, SNA recibe la información de los diferentes órganos y del medio interno, para actuar sobre la frecuencia cardíaca y respiratoria, la contracción y dilatación de vasos sanguíneos, la digestión, la salivación, el sudor, la contracción y relajación del músculo liso en varios órganos, acomodación visual, dilatación de la pupila, secreción de glándulas exocrinas y endocrinas, la micción y la excitación sexual (Braunwald, 2009). El SNA se divide funcionalmente en el sistema simpático (SNS) y parasimpático (PNS). El SNS está implicado en actividades que requieren gasto de energía, y es el que prepara al individuo para reaccionar ante una situación de estrés. El PNS está encargado de almacenar, conservar la energía y es el que mantiene al individuo en situaciones normales que no implican peligro o estrés, vale decir actúa de manera opuesta al SNS.

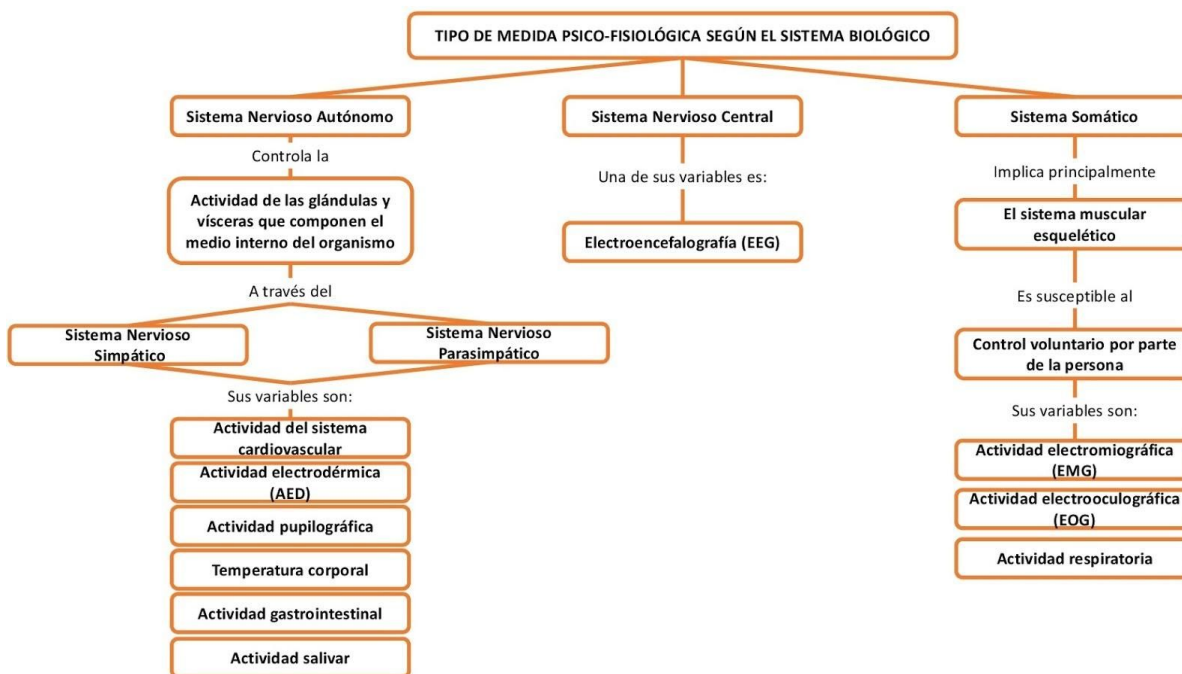


Figura 1: Mapa conceptual de medidas psicofisiológicas según su sistema nervioso.

Entre las señales psicofisiológicas con mayor poder predictivo de fatiga, se encuentran la variación del ritmo cardíaco (HRV), actividad electrodermal (EDA), actividad cerebral, dilatación pupilar, temperatura en la piel y la voz. A continuación, se explica cómo las señales psicofisiológicas mencionadas se ven afectadas ante una situación de actividad física o mental:

- **Ritmo cardiaco (HR):** el estrés agudo hace que el corazón se contraiga con gran fuerza y frecuencia. Con estrés crónico, la masa del corazón se incrementa para proporcionar al cuerpo una mayor respuesta a los factores estresantes (Devereux, et al., 2000). Una disminución en la amplitud de la onda medida por un electrocardiograma es un indicador de estrés en individuos sanos (Bonoris, et al., 1978). Se sabe que bandas de baja frecuencia (LF, 0 a 0,15 Hz) y alta frecuencia (HF, 0,16 a 0,4 Hz) del HRV reflejan actividades de los sistemas nerviosos simpático (SNS) y parasimpático (PNS), ya que la actividad del SNS está asociada con la banda LF y el PNS con la banda HF, con lo que se puede concluir que el nivel de estrés incrementa cuando aumenta la actividad LF y disminuye la actividad HF del HRV (J. Zhai & A. Barreto, 2006).
- **Actividad electrodermal (EDA):** es una medida del flujo de electricidad a través de la piel de un individuo. Cuando el individuo está bajo estrés, la conductancia de la piel aumenta (W. Liao, et al., 2005) debido al aumento de la humedad en su superficie, lo que incrementa el flujo de electricidad. Por el contrario, la conductancia de la piel se reduce cuando el individuo disminuye su estrés.
- **Actividad cerebral:** La actividad neuronal en el cerebro produce señales eléctricas que se pueden capturar usando electroencefalografía (EEG). La EEG registra la forma de las ondas eléctricas en el cuero cabelludo formado por potenciales eléctricos de acción durante excitaciones sinápticas e inhibiciones de dendritas. Los potenciales en el cuero cabelludo varían entre 20 a 100 μ V (D. Novák, et al., 2004). La actividad en el hemisferio derecho del cerebro domina durante las emociones negativas, mientras que la actividad en el hemisferio izquierdo domina durante las emociones positivas (R. Horlings, et al., 2008), lo que sugiere un área para la detección del estrés.
- **Dilatación pupilar:** Si el diámetro de la pupila de un individuo aumenta, entonces sugiere que el individuo está posiblemente en un estado estresado (W. Liao, et al., 2005). Sin embargo, tanto los estímulos negativos como los positivos pueden hacer que los diámetros de las pupilas aumenten y se debe tener cuidado con aislar el estímulo lumínico.
- **Voz:** son los componentes no verbales de la voz los que reflejan el estrés. Las características de voz para los modelos de estrés incluyen sonoridad, frecuencia fundamental, relación de frecuencia de fluctuación y energía (Rothkrantz, et al., 2004). Las emociones relacionadas con el estrés se midieron en el habla extrayendo características que incluyen calidad de voz, tono, duración, intensidad, entre otras (Scherer, et al., 2008).

- **Electromiografía (EMG):** Es el registro de la actividad eléctrica producida por la contracción muscular. Cuando la neurona motora envía una señal al músculo para que se contraiga, la actividad eléctrica es producida por las células musculares. El dispositivo para cuantificar EMG registra la actividad agregada de muchas células motoras que disparan al mismo tiempo. Para la detección del estrés, se suele colocar electrodos EMG en el músculo trapecio, que se encuentra en el hombro. Se ha estudiado que el estrés mental causa un aumento en la actividad de EMG siguiendo el mismo patrón de la respuesta de la presión arterial y el estrés autoinformado. El estrés psicológico desempeña un papel en los trastornos musculoesqueléticos al aumentar la tensión muscular tanto en situaciones de baja carga como en ausencia de carga física; además, la tensión muscular causada por el estrés se acentúa durante la ejecución de una carga física.

Si bien las medidas psicofisiológicas constituyen un método más insesgado que la sola aplicación de un cuestionario autoinformado, aun así poseen cierto grado de subjetividad, ya que no todas las personas reaccionan de la misma manera ante un mismo estímulo, debido a que la actividad de su fisiología depende de su estado de salud. Esto sugiere la utilización de técnicas cuantitativas para la normalización de las mediciones y el filtrado de esta fuente de incertidumbre.

Análisis estadístico y aprendizaje de máquina

El análisis de las señales psicofisiológicas buscan por un lado identificar y describir eventos de fatiga y por otro lado clasificarlos en intensidad y buscar patrones de ocurrencia. Para cada uno de estos objetivos la literatura propone el uso de distintas familia de análisis: Análisis estadístico y Aprendizaje de máquina.

La técnica estadísticas destacan a la hora de caracterizar y describir un fenómeno gracias a que se pueden usar para probar significancia estadística en el efecto de una variable sobre otra. Aunque han perdido terreno en comparación con los métodos de aprendizaje de máquinas, poseen la ventaja de necesitar un menor número de datos y de poder explicar el fenómeno observado. Algunas de las técnica utilizadas en la literatura son el test T de Student (Kim & Cha, 2015) y el análisis de varianza (ANOVA) (Wegge, et al., 2007).

Entre las técnicas de aprendizaje de máquinas para clasificar niveles de fatiga entre "alto", "bajo" y "estado de relajación" a partir de señales psicofisiológicas, destacan los árboles de decisión (Aryal., et al., 2017), las redes neurales artificiales (Wilson & Russell, 2003) y las Support Vector Machines (SVM) (Zhai & Barreto, 2006). También existen técnicas basadas en el modelamiento probabilístico, que se pueden utilizar para medir estrés, donde destacan las Dynamic Bayesian Networks (DBN) (W. Liao, et al., 2005) y los Hidden Markov Models (HMM) para reconocimiento de estrés utilizando registro de la voz (M. Rahurkar, et al., 2002). Se observa que mientras más complejo

sea el modelo matemático, mejor es la precisión en la clasificación pero más difícil es la interpretación de los resultados. Por otro lado, si bien estos modelos logran clasificar y predecir, no están diseñados para concluir significancia estadística a través de sus resultados. Además, los modelos de aprendizaje de máquina, son intensivos en la cantidad de datos necesarios para lograr buenos niveles de precisión, el número exacto de la cantidad de datos necesarios varía por cada modelo y aplicación.

Previo al análisis de datos de señales psicofisiológicas es necesario realizar un pre-procesamiento para lidiar con incertidumbres provenientes de las variaciones intra e inter individuos de los datos auto reportados, variaciones individuales en las señales psicofisiológicas y quitar valores anómalos, ruido y suavizar las señales. Para esto se han utilizado técnicas de filtrado y de separación en bandas de frecuencia (uso de Wavelet en Daubechies, 1990), Deep Belief Networks (DBNs) para eliminar datos erróneos (Son, et al., 2018) y filtros difusos y de Finite-Mixture Models para lograr hacer una segmentación de los niveles de fatiga de una forma menos sesgada (Kumar, 2008).

Conclusiones del estado del arte

La utilización del enfoque psicofisiológico para la medición de estrés presenta ventajas con respecto a otros enfoques al lograr mediciones poco invasivas, de gran precisión, cuantificable y en tiempo real. Las señales con mejores propiedades son el ritmo cardíaco, la actividad electrodermal y características de la voz. Otras señales como la actividad cerebral y dilatación pupilar son descriptores significativo de fatiga pero no recomendables para el estudio actual dado su alto costo de implementación y complejidad de análisis. La electromiografía muestra ser una herramienta especialmente relevante en la detección de la fatiga física, a pesar de la invasividad de su uso.

Las metodologías experimentales utilizadas presentan 3 etapas en la medición de señales psicofisiológicas: medición del "estado basal", "en trabajo" y "en recuperación". Adicionalmente a las mediciones psicofisiológicas se agregan métricas auto informadas durante la etapa de "en trabajo". Estas métricas buscan validar lo medido psicofisiológicamente o cuantificar el estrés (i.e. escala del 1 al 5). Se sugiere también medir y/o controlar variables de contexto y el estado de los participantes para aislar el efecto de la labor en el estrés medido. En cuanto a la muestra experimental se observa una gran variabilidad tanto en el tamaño como en la demografía de los sujetos experimentales. Sin embargo el mínimo número de participantes en un estudio es de 9, marcando así una cota inferior para el tamaño muestral.

El análisis de las señales se realiza principalmente con técnicas de aprendizaje de máquinas. Métodos tales como filtros difusos, mezcla finita de modelos y Deep Belief Networks son utilizados

para eliminar sesgos, incertidumbres e inconsistencias en los datos. En especial para este proyecto los modelos de clasificación muestran excelentes resultados en la detección de estrés y fatiga. Sin embargo estos métodos sólo permiten clasificar dos estados ("relajado", "estresado") lo que es insuficiente para nuestras necesidades. Esto último abre una interesante arista de investigación.

2.- METODOLOGÍA PARA LA PREDICCIÓN DE FATIGA

A continuación, se presenta la metodología para la caracterización de una tarea en función de la fatiga que produce durante su ejecución. La figura 2 muestra la metodología propuesta en la que se definen cuatro etapas: i) Definición de la tarea, ii) Definición del experimento, iii) Preparación de datos y iv) Análisis de la tarea.

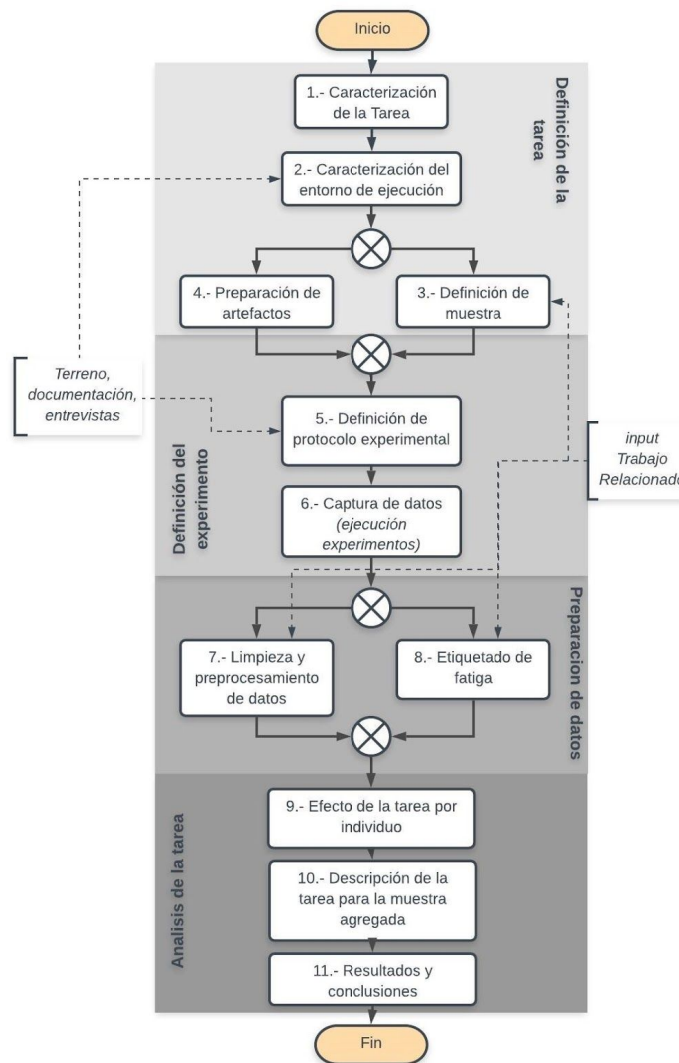


Figura 2: Metodología para la caracterización de tareas laborales en función de la fatiga que generan sobre los trabajadores, utilizando señales psicofisiológicas.

A.- DEFINICIÓN DE LA TAREA

El objetivo de esta etapa es la caracterización de la tarea en término de las actividades que la componen, la estructura de turnos, el entorno en el que se desarrolla, los trabajadores que las practican y la definición de la muestra a tomar. Todos estos factores representan insumos para la siguiente etapa de definición del experimento para la captura de datos y la evaluación de la fatiga. Como se muestra en la figura 1 esta etapa se compone de 4 subtareas, las que se describen a continuación: 1) Caracterización de la tarea, 2) Caracterización del entorno de ejecución, 3) Definición de la muestra y 4) Preparación de artefactos.

1.- Caracterización de la tarea

La caracterización de la tarea resulta fundamental para realizar la toma de datos de la manera más realista posible. Es importante realizar la descripción de la tarea a través de la descripción de las actividades a realizar y a la forma en la que éstas se realizan. La descripción de esta subtarea se esquematiza en la Tabla II. A continuación, se describen los puntos mínimos a incluir para la descripción de la tarea: i) Monotonía de la actividad, ii) Repetitividad de la actividad, iii) Continuidad en la ejecución de la actividad y iv) Métricas de rendimiento.

Tarea	Actividades	Monotonía	Repetitividad	Continuidad
Nombre de la tarea	Actividad 1 Actividad 2	(no) Monotona	(no) Repetitiva	(no) Continua

Tabla II: Caracterización de la Actividad

i) Monotonía de la actividad: En este punto se busca describir la actividad en función de las actividades relevantes (o hitos) a realizar durante la actividad. Esto se realiza mediante visitas técnicas al lugar del trabajo, consulta a los trabajadores, prevencionistas de riesgo y/o ergónomos.

Las actividades pueden ser Monótonas o no Monótonas. Mientras una actividad monótona presenta actividades pequeñas con solo un hito en su ejecución una actividad no monótona presenta distintos grupos de actividades que buscan cumplir con un número de hitos necesarios para completar una actividad.

ii) Repetitividad de la actividad: En este punto resulta relevante describir si la tarea sigue un patrón recurrente en su ejecución o no. Ejemplo de tareas repetitivas son digitación o reposición de góndolas. Conducción es ejemplo de una tareas no repetitiva.

iii) Continuidad en la ejecución de la tarea: La continuidad de la ejecución hace referencia a la existencia de descansos durante la ejecución de la tarea. Actividades tales como digitación o reposición de góndolas no incluyen descansos en su ejecución mientras que actividades como atención al cliente presentan momentos donde el trabajador no se encuentra realizando su tarea principal.

iv) Medición de Rendimiento: Basado en la descripción de la tarea se deben definir las métricas de rendimiento que mejor se ajusten a la tarea en estudio. No obstante, la métrica específica deberá siempre considerarse métricas de eficiencia, eficacia y de productividad (Tabla III). El análisis de estas métricas permite la evaluación de la degeneración de la ejecución, lo que se traduce en la métrica de fatiga de rendimiento (Punto C.8 Etiquetado de Fatiga).

Error Total	Métrica de eficacia	Métrica de eficiencia	Métrica de productividad
Promedio simple de las tres métricas definidas	Número de errores	Frecuencia de errores	Tiempo promedio de ejecución

Tabla III: Métricas de rendimiento para la actividad en estudio

2.- Caracterización del entorno de ejecución

El entorno de ejecución de la tarea define una serie de restricciones tanto para la realización de la tarea como para la medición de factores externos que afecten al trabajador durante su labor. En general, se definen dos ámbitos del entorno de ejecución: i) entorno directo y ii) entorno indirecto.

- **Entorno directo:** Entorno más próximo al trabajador, en el que se encuentran sus herramientas de trabajo y que es controlable por él mismo o de alguna forma directa.
- **Entorno indirecto:** entorno que no es posible controlar directamente pero que afecta las tareas del trabajador. No necesariamente existen ambos entornos para toda tarea, puede haber tareas que tengan solo uno de ellos. El entorno indirecto incluye condiciones físicas del entorno al momento de la ejecución, características demográficas de los sujetos y condiciones de inicio del experimento.

3.- Definición de la muestra

Las muestras experimentales para el análisis a realizar posee dos dimensiones principales: i) Número de usuarios y ii) Número de unidades de observación. Mientras la primera dimensión entrega validez estadística a los resultados la segunda permite la aplicación de herramientas estadísticas y de aprendizaje de máquinas.

i) Número de sujetos experimentales: Esta dimensión hace referencia a la cantidad de sujetos que se someterán a las actividades experimentales. En la literatura se estima que un número de alrededor de 20 sujetos entregan validez estadística para experimentos que analizan fatiga y que consideran señales psicofisiológicas y/o cuestionarios de fatiga (Patel, 2011; Ma, 2018; Fu, 2016; Lin, 2014; Chuang, 2018; Zhang, 2014; Aryal, 2017). El mínimo de sujetos experimentales encontrados en la revisión bibliográfica corresponde a 12 (Patel, 2011; Fu, 2016). De existir un caso donde este número no es posible de obtener se deberá trabajar con el número existente pero se deberá someter a los sujetos al experimento en múltiples ocasiones (muestras repetitivas) para lograr el número deseado de 20 muestras. Cabe mencionar que mientras más sujetos y mientras más cercana sea la muestra al universo de sujetos más generalizables serán los resultados. Los sujetos experimentales deberán ser seleccionados considerando los criterios de inclusión y exclusión correspondientes a la actividad a estudiar.

ii) Número de unidades de observación: El número de unidades de observación corresponde al número de muestras de señales psicofisiológicas a estudiar. Por sujeto de experimentación existen múltiples unidades de observación. Las unidades de observación son el resultado de la división temporal de las señales capturadas. Estas unidades de observación o ventanas de tiempo (ej. promedio de una señal cada 5 segundos) son el objeto de análisis y su número deberá permitir la aplicación de las herramientas de aprendizaje de máquinas. En general el número agregado de unidades de observación deberá ser del orden de 10.000 muestras considerando todos los sujetos experimentales. Utilizando este número es que se deberá fijar el tiempo de ejecución de la actividad por sujeto.

4.- Preparación de artefactos

En esta etapa se deberá definir y fijar los artefactos para las distintas mediciones. Esto incluye la definición de los mecanismos de mediciones i) psicofisiológicas, ii) contextuales y iii) formularios legales. Cada una de ellas se explican a continuación.

i) Medidas psicofisiológicas: En esta metodología se contempla la medición del pulso cardíaco y su variación, temperatura superficial de la piel y actividad electrodermal. Adicionalmente se mide la inercia (aceleración y orientación) de la herramienta de sensado por requerimientos técnicos de filtrado de datos (Sección D.7 “Limpieza y procesamiento de datos”).

Dentro de esta metodología se propone el uso del sensor BioMonitor V2¹. Esta pulsera (Figura 3) es una plataforma de sensado inalámbrico que realiza múltiples mediciones de señales psicofisiológicas, las que envía a un receptor central mediante bluetooth 4.0 (para detalles sobre la implementación del sistema de recepción ver Anexo II: Docking Station). La Tabla IV describe las señales y características con las que cuenta la pulsera.

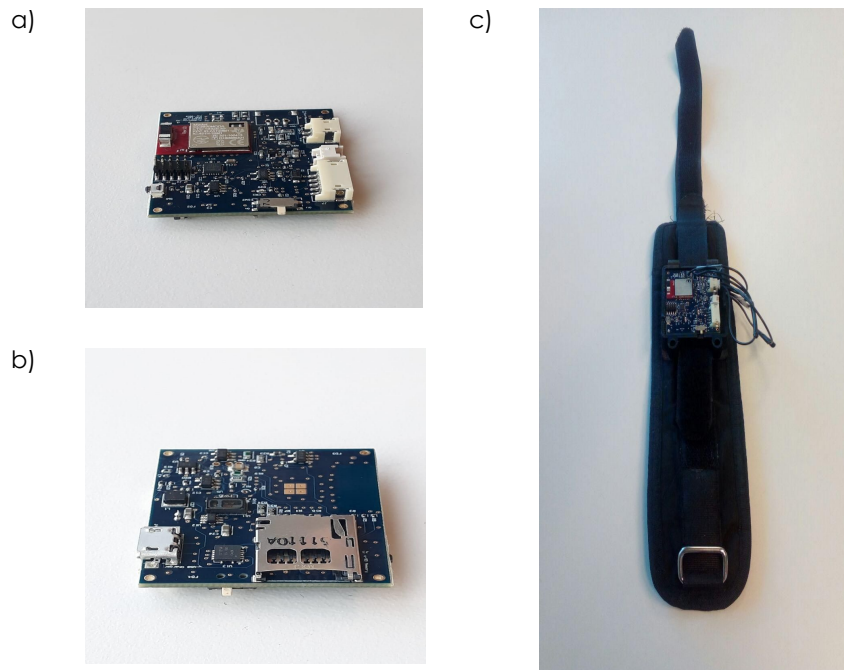


Figura 3: Plataforma de sensado Biomonitor V2.

Sensor	Descripción sensores incluidos	Lecturas por segundo
Temperatura (GA10K3A1AM)	Sensor análogo de temperatura (C°), rango de 31°C y 44°C	entre 1 y 20
EDA (GARY)	Sensor análogo de impedancia de la piel (mAh)	entre 1 y 100

¹ Es posible el uso de otras plataformas de sensado sin embargo es necesario cumplir con las especificaciones técnicas descritas en la tabla IV

PPG (Max30101)	Sensor análogo que mide la absorción de haz de luces en la piel	50 ó 100
Inercial MPU9250 (acelerómetro y giroscopio)	Sensor análogo de aceleración (g) y velocidad angular (rad°)	entre 5 y 100

Tabla IV: Descripción de los sistemas de sensado de señales psicofisiológicas.

ii) Medidas Contextuales: Las medidas contextuales entregan información directa e indirecta del medio donde se desenvuelve el sujeto de prueba y permiten estimar sus efectos sobre la fatiga. Se distinguen dos dimensiones del contexto principales: i) Contexto de entorno y ii) Características del sujeto experimental. Las métricas específicas utilizadas para la caracterización de cada contexto se encuentra en la tabla V.

Contexto	Descripción sensores incluidos
Entorno	Plataforma de monitoreo contextual: CO2, Temperatura ambiental, Ruido ambiental, Movimiento, Humedad (Figura 7)
Características del sujeto	Formulario autoinformado: Nombre, Edad, Género, Estado Civil, Hijos, Educación, Profesión/oficio, Tipo de Contrato, Período en la empresa, Experiencia Laboral, Realiza actividad física, ¿Durmió bien anoche?, Horario en el que durmió, Cantidad de horas de sueño.

Tabla V: Medidas contextuales consideradas.

- Contexto de entorno:** Las condiciones de entorno afectan de manera importante tanto el rendimiento como la fatiga sufrida por el trabajador. Dentro de las variables relevantes descritas en trabajos relacionados en casos de medición de fatiga se encuentran: concentración de CO2, temperatura ambiental en el lugar de desarrollo de la tarea, nivel de ruido y luminosidad (Fu, 2016).
Dentro de esta metodología se contempla el uso de una unidad de sensado de elaboración propia (figura 4) que mide: Ruido ambiente, concentración de CO2, movimientos Inerciales (acelerómetro y giroscopio), temperatura ambiental y humedad.
- Características del sujeto experimental:** A través de un cuestionario se registran características demográficas y experiencia laboral. Identificar la experiencia propia y realidad personal, que puedan afectar su rendimiento, es un complemento a las variables del contexto directo de la tarea. Por ejemplo, factores de salud, cantidad y calidad de sueño, ciclo circadiano, la seguridad de la fuente laboral y problemas

familiares, tiene un impacto en el nivel de estrés y fatiga de una persona. Esta información se obtiene a través de cuestionarios.

iii) Formularios legales: Estos varían según la actividad y el entorno de la empresa u organización. No obstante, en la metodología se considera el uso de un formulario de consentimiento informado para cada sujeto experimental. Este consentimiento debe informar al sujeto sobre las señales que se capturarán, para qué se capturarán y quién es el responsable de almacenarlas. Deberá contar con firma, nombre y aceptación del usuario.

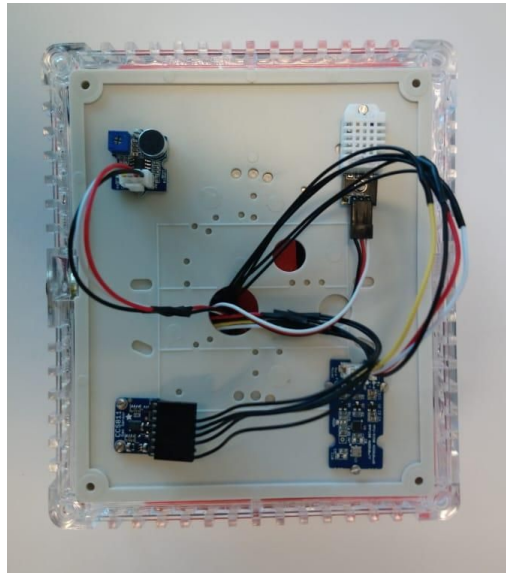


Figura 4: Plataforma de sensado de contexto

B.- DEFINICIÓN DEL EXPERIMENTO

De acuerdo con la metodología en esta sección se diseña y ejecuta el experimento. Concretamente se define el protocolo experimental, se ejecutan los experimento y se captura los datos experimentales.

5.- Definición del protocolo experimental

La formulación del protocolo experimental requiere la definición detallada de las actividades a realizar antes, durante y después del experimento. Aunque los pasos son específicos a la actividad a analizar, al tipo de sujetos experimentales y a las condiciones del experimento, existen actividades transversales que permiten generar un protocolo experimental dentro de la metodología. A

continuación, se presenta el mínimo de las actividades a considerar en el protocolo experimental para cada actividad (Tabla VI).

Antes de experimento	Durante el experimento	Después del experimento
1.- Verificación de los instrumentos, aplicaciones y área de experimentación. 2.- Recepción, registro de información del participante y descripción general del experimento. 3.- Instalación de sensores y adaptación. 4.- Explicación de subtareas o actividades complementarias. 5.- Registro de línea base.	1.- Explicación detallada de la ejecución de la tarea. 1.- Inicio y término de la tarea. 2.- Supervisión de la ejecución y monitoreo de los sensores. 3.- Toma de notas del experimentador.	1.- Registro de línea base final. 2.- Detención de captura de datos y desmontaje de sensores. 3.- Registro de incidentes reportados por el sujeto. 4.- Descarga y almacenamiento de datos.

Tabla VI: Actividades mínimas para la definición del protocolo experimental.

6.- Ejecución del protocolo experimental

Esta etapa corresponde a la ejecución del experimento y la toma de datos. Es relevante informar el resumen final de la toma de datos para considerar cualquier anomalía en las etapas de procesamiento de datos. El reporte de esta etapa consiste en tres entregables principales: i) Datos crudos de los sensores y ii) Datos de herramientas y artefactos aplicados a los sujetos de experimentación, por ejemplo formularios completados y iii) Anotaciones del experimentador. Estos datos se deberán entregar en formato digital y siguiendo el formato especificado en el protocolo experimental.

C.- PREPARACIÓN DE DATOS

Antes de utilizar estos datos para la caracterización de la tarea y la descripción de la fatiga es necesario examinar estos datos y prepararlos para la aplicación de técnicas estadísticas y de aprendizaje de máquinas. Lo anterior para eliminar cualquier contaminación de las señales capturadas por los sensores y para completar los datos de sensores con los otros datos capturados (sensores ambientales, notas del experimentador, etc). En particular se presentan a continuación las actividades de i) Limpieza y procesamiento de datos y ii) Etiquetado de fatiga.

7.- Limpieza y procesamiento de datos

Las señales capturadas por los sensores representan las bioseñales más una serie de otras señales provenientes del ambiente, del movimiento de los sujetos de prueba y de los propios mecanismos de los sensores. Dado a lo anterior resulta fundamental preparar los datos para asegurar la correcta obtención de información a partir de estos.

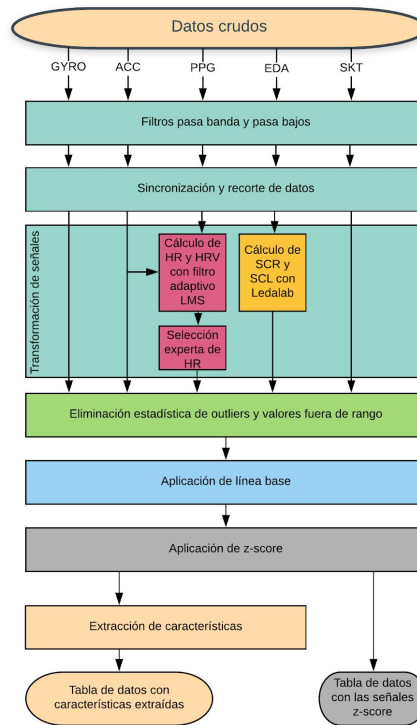


Figura 5: Proceso de filtrado y limpieza de datos psicofisiológicos

Como se muestra en la Figura 5, esta etapa se compone de un número de etapas en las que los datos crudos capturados por los sensores psicofisiológicos son procesados hasta obtener 3 entregables que se utilizan luego en el procesamiento de datos: Tabla de datos de las señales psicofisiológicas, Tabla de señales con línea base y Tabla de datos de las señales con línea base y z-score.

A continuación, se presentan las actividades relacionadas con la limpieza, preprocesamiento y preparación de datos de sensores para su posterior análisis (Figura 4): i) Filtrado pasa bandas y pasa bajos, ii) Sincronización y recorte de datos, iii) Transformación de señales psicofisiológicas, iv) Eliminación de outliers y valores fuera de rango, v) Aplicación de línea base y vi) Aplicación de z-score.

i) Filtrado pasa bandas y pasa bajos: De acuerdo a la literatura es fundamental manejar el ruido de las señales por movimiento (Haapalainen et al., 2010). Esta se produce cuando los sensores son sometidos a movimientos durante el sensado. Ejemplos de esto son movimientos al correr, saltar o movimientos debido un mal ajuste de los sensores. El primer cuidado a considerar es ajustar correctamente los sensores al cuerpo, disminuyendo así el ruido generado por la actividad del sujeto de prueba. El segundo paso es la aplicación de algoritmos para la corrección de las señales capturadas.

Como primer paso se aplican filtros pasa bajos y pasa bandas a las señales obtenidas desde los datos crudos. Con estos filtros se busca eliminar componentes no propias de las señales que se encuentren en rangos de alta o baja frecuencia (Figura 6). Cada señal es filtrada con rangos propios que se detallan a continuación:

- ACC(X,Y,Z): Filtro pasa banda con frecuencia de corte entre 0,5 y 1 Hz.
- PPG: Filtro pasa banda entre 0.4 y 8 Hz.
- EDA: Filtro pasa bajos con ventana de Hanning con frecuencia de corte de 5 Hz.
- SKT: Filtro pasa bajos con frecuencia de corte de 0,5 Hz.

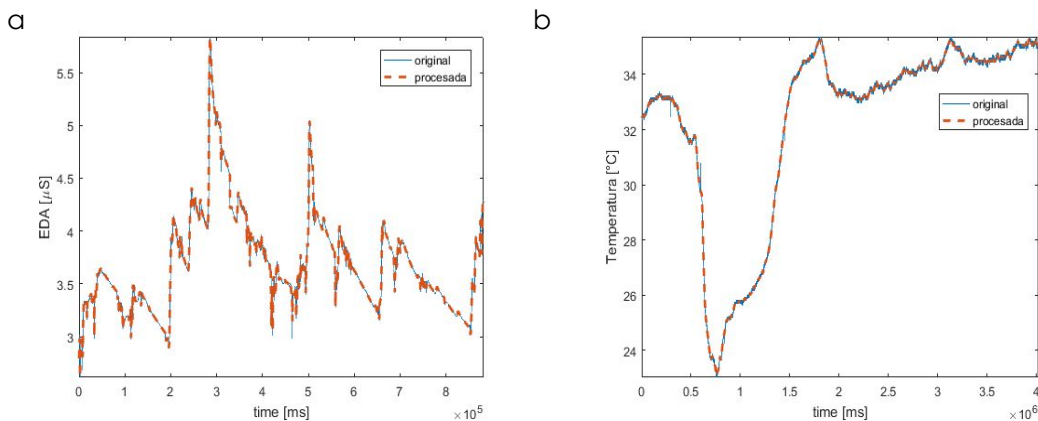


Figura 6: Preprocesamiento de señales de EDA (a) y Temperatura (b).

ii) Sincronización y recorte de datos: Una vez filtrados los datos se deben sincronizar temporalmente e identificar dentro de la ejecución del experimento. Como se establece en el protocolo experimental, dentro de toda la muestra de datos existen tres etapas principales: línea base (BASELINE), Ejecución de la tarea (TASK) y línea base final (RECOVERY).

En esta etapa se sincronizan temporalmente las distintas señales psicofisiológicas en una tabla única, generando así un vector de mediciones por cada instante en que los sensores capturaron datos. Posteriormente, se agregan a esta tabla única los datos contextuales obtenidos por el sensor de contexto y las respuestas de los formularios.

iii) Transformación de señales: No todas las señales capturadas representan directamente señales psicofisiológicas. Si bien en el caso de la temperatura la señal capturada corresponde a la temperatura directamente, en el caso del PPG y de EDA la señal capturada no es la señal psicofisiológica directamente. El objetivo de esta etapa es la de transformar las señales obtenidas en las señales psicofisiológicas necesarias para el análisis posterior.

A continuación se explica la metodología para la transformación requerida sobre las señales de PPG y EDA (estado del arte hasta el momento).

Transformación del PPG: Cálculo de ritmo cardíaco (HR) y variabilidad del ritmo cardíaco (HRV)

El algoritmo para calcular HR y HRV utiliza la señal del acelerómetro triaxial para detectar los artefactos de movimiento que introducen ruido a la señal PPG. Lo primero es realizar un resamplado de las señales de ACC y PPG a 125 Hz, valor en el que está seteado el algoritmo.

Mediante el método "*Singular Value Descomposición*" (SVD) aplicado a la señal simultánea de aceleración se extraen las componentes de referencia para un filtro adaptivo que se aplica al PPG que elimina la mayor parte del ruido causado por el movimiento (Mashhadi, 2016; Zhang, 2015). Con esta señal procesada se calculan el HR y HRV.

Para calcular el HR se aplican 2 métodos: i) análisis espectral de la señal PPG limpia (análisis en el plano de frecuencia) y ii) análisis en el espectro temporal. El cálculo de HRV también se realiza directamente en el plano temporal. Todo esto se realiza en ventanas de 8 segundos con un deslize de 2 segundos (Mashhadi, 2016).

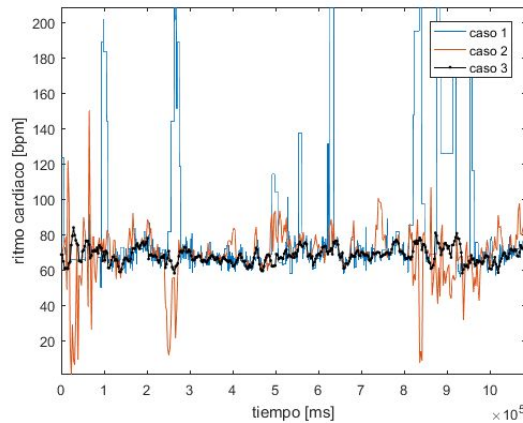


Figura 7: Preprocesamiento de señales de HR en 3 casos: Caso 1: Procesamiento sin considerar el filtraje de movimiento. Caso 2: Procesamiento con filtrado básico de movimiento. Caso 3: Procesamiento considerando acelerometría (método antes expuesto).

Cabe indicar lo relevante del ruido por movimiento como un factor crítico en el procesamiento de señales psicofisiológicas. Además, el uso del acelerómetro como medida de referencia para el filtrado del movimiento en las señales PPG brinda buenos resultados para calcular HR y HRV. Esto resultados quedan de manifiesto en la Figura 7.

Transformación del EDA: Cálculo de Nivel de Conductancia de la Piel (SCL) y Respuesta de la Conductancia de la Piel (SCR)

El cálculo de estas componentes se realiza mediante la extracción de las componentes fásica (SCL) y tónicas (SCR) de la señal limpia de EDA. Para esto se realiza un análisis de descomposición continua (CDA, *Continuous Decomposition Analysis*). (Benedek, M. & Kaernbach, C. (2010)). Este análisis arroja las señales en sus componentes SCL y SCR.

Discriminación experta de las señales de HR

La discriminación de las señales de HR se realiza por un experto que considera valores dentro de rango entre otros parámetros para escoger la mejor de las dos señales.

iv) Eliminación de outliers y valores fuera de rango: Luego de los pasos anteriores aún es necesario refinar los resultados de las señales crudas. En esta etapa se busca eliminar datos que se encuentren estadísticamente fuera de rango y que estén también fuera de rangos fisiológicos que no hayan sido detectados o eliminados anteriormente. Esto es realizado

también por un experto en señales. El resultado final de esta etapa es un dataframe con las señales psicofisiológicas en sus rangos originales.

v) Aplicación de línea base: En esta etapa se resta la línea base a cada señal de forma de eliminar efectos individuales (características psicológicas, físicas o demográficas de los individuos) y poder comparar a cada sujeto en cada una de sus etapas de experimentación (TASK, RECOVERY). El resultado final de esta etapa es un dataframe con las señales psicofisiológicas con rangos comparables en cada etapa del experimento para cada sujeto.

vi) Aplicación z-score: En esta etapa se estandarizan los rangos de las mediciones de cada individuo en consideración de la muestra. De esta forma se obtienen señales psicofisiológicas en rangos que permiten la comparación entre-sujetos (Guyon, 2016). El resultado final de esta etapa es un dataframe con las señales psicofisiológicas estandarizadas para cada sujeto.

vii) Extracción de características: Una vez limpias y calculadas todas las señales psicofisiológicas se procede a calcular el conjunto de variables psicofisiológicas final utilizado para el análisis en el punto D. La figura 7 muestra estas variables, las que se conocen con el nombre de características (*features*) de las señales (Guyon, 2016).

El primer paso en la extracción de características es definir el tamaño de la ventana de tiempo en el cual se observarán los datos. Para esto se debe considerar la descripción de la tareas (punto 1.A.2) en términos de su repetitividad y monotonía. En general se recomienda una ventana de 5 minutos a excepción de que la labor tome más de 5 minutos en ser realizada (von Rosenberg et al., 2017). En este último caso se recomienda aumentar la ventana de forma que esta considere a lo menos una ejecución de la tarea. Para aumentar la cantidad de muestras se recomienda que exista un traslape entre las ventana realizando un deslice de 10 segundo.

Luego para cada individuo y para cada una de sus señales se calcula cada una de las características que se presentan en figura 8. Finalmente, se realiza una selección de características con el algoritmo Boruta, el que entregará un resultado específico para cada actividad (Kursa, et al., 2010).

PPG	meanHR	Promedio del ritmo cardíaco
	stdHR	Desviación estándar del ritmo cardíaco
	rangeHR	rango del ritmo cardíaco
	meanHRV	Promedio de la variabilidad del ritmo cardíaco
	stdHRV	desviación estándar de la variabilidad del ritmo cardíaco
	rmssdHRV	Error cuadrático medio de la variabilidad del ritmo cardíaco
	VLF	Potencia de la componente de frecuencia del HRV entre 0,003-0,04 Hz
	LF	Potencia de la componente de frecuencia del HRV entre 0,04-0,15 Hz
	HF	Potencia de la componente de frecuencia del HRV entre 0,15-0,4 Hz
	TP	Potencia total de la componente de frecuencia del HRV
	ratioLFHF	Tasa entre LF y HF
EDA	meanEDA	Promedio de la actividad electrodérmica
	stdEDA	Desviación estándar de la actividad electrodérmica
	meanSCL	Promedio del nivel de conductancia de la piel
	meanSCR	Promedio de la respuesta de conductancia de la piel
	rangeSCL	Rango del nivel de conductancia de la piel
	maxSCR	Valor máximo de la respuesta de conductancia de la piel
	NpeakSCR	Número de peaks de la respuesta de conductancia de la piel
SKT	meanSKT	Promedio de la temperatura corporal
	stdSKT	Desviación estándar de la temperatura corporal
	rangeSKT	Rango de la temperatura corporal

Figura 8: Características extraídas para cada muestra.

8.- Etiquetado de fatiga

En esta etapa se completa la tabla única de datos con la información de fatiga capturada durante la ejecución de la tarea agregando los eventos de fatiga capturados. De esta forma se logra etiquetar momentos de fatiga, los que luego son cotejados con las señales capturadas psicofisiológicas. Para lograr esto es necesario definir una ventana de observación de fatiga y luego etiquetar esas ventanas, de ser posible, en función de las tres métricas de fatiga definidas: Fatiga autoreportada, Fatiga de rendimiento y Fatiga observada. A continuación, se describen estas y su forma de medición.

i) Etiquetado de fatiga para las ventanas de observación:

- Fatiga autoreportada:** Corresponde a la fatiga que los mismos sujetos de prueba reportan al ser consultados. Para el desarrollo de la presente metodología se utiliza el cuestionario *Karolinska Sleepness Scale* (KSS) que mide el nivel subjetivo de somnolencia autorreportado durante un período específico (10 minutos en nuestra metodología) (Akerstedt, T., & Gillberg, M. (1990)). Este consiste en una pregunta '¿Cómo te sientes?' respecto a una explicación previa de la escala de somnolencia consistente en una escala de Likert de 1 a 9 (Tabla VII), con los siguientes valores:

Nivel	Descripción
1	Extremadamente Alerta
2	Muy Alerta
3	Alerta
4	Algo Alerta
5	Ni Alerta ni Somnoliento
6	Alguna señal de Somnoliento
7	Somnoliento, sin esfuerzo para mantenerse despierto
8	Somnoliento, poco esfuerzo para mantenerse despierto
9	Muy Somnoliento, alto esfuerzo para mantenerse despierto, lucha con el sueño

Tabla VII: Niveles de somnolencia del test KSS

El KSS se ha utilizado en estudios de trabajos por turnos, desfases horarios, en conducción, entre otros, tanto en hombres como mujeres. Además, en un estudio realizado por Kaida et al. (2006), se estudió la validez del KSS y se encontró que existe una alta correlación con el EEG y las variables de comportamiento. Para evaluar la confiabilidad de este test se calcula el Alfa de Cronbach que debe ser mayor a 0,8 para considerarlo altamente confiable

- **Fatiga de rendimiento:** Corresponde a la fatiga que se refleja en la disminución de eficiencia y eficacia en la ejecución de la tarea. Para esto se debe analizar el resultado de la ejecución del trabajo en función de las medidas de eficiencia y eficacia definidas en el punto A.1. "Caracterización de la Tarea". El indicador de fatiga de rendimiento final es una ponderación de las métricas definidas en A.1, obteniéndose así solo un valor de fatiga de rendimiento. Un ejemplo de test de rendimiento se presenta a continuación.
 - **CTP:** Las pruebas de rendimiento continuo (CPT) son usados frecuentemente para obtener información cuantitativa de la habilidad de un individuo para mantener la atención en el tiempo (Ricció, 2002). En general, un CPT mide la atención selectiva y sostenida de una persona, donde la atención sostenida a la capacidad de mantener el foco en un estímulo continuo, y la selectiva se

refiere a la capacidad de mantener el foco en estímulos seleccionados e ignorar las distracciones.

- **N-back:** es un CPT para medir la retención de corto plazo (Kirchner, 1958), con el paso del tiempo existen varias versiones de N-back, que adaptadas pueden ser aplicadas con seguridad durante la tarea de conducción, tanto en simuladores como en ruta (Mehler, 2010). Este funciona presentando una serie de elementos (número, letras, etc) y pidiendo al sujeto de prueba que repita el N-anterior al elemento actual.
- **Fatiga observada:** Corresponde a gestos de fatiga observados por el experimentador. Este cálculo se realiza en función del reporte de experimentación y se asignan puntajes de acuerdo a la tabla VIII. La puntuación se realiza sumando el puntaje de todos los incidentes listados en la tabla VIII para cada ventana de observación (Sahayadhas,2012).

ACCIÓN	PUNTAJE	DESCRIPCIÓN
SUSPIRO LEVE	1	EXHALAR AIRE POR BOCA Y/O NARIZ, LEVE
SUSPIRO	2	EXHALAR AIRE POR BOCA Y/O NARIZ, NORMAL
SUSPIRO PROFUNDO	3	EXHALAR AIRE POR BOCA Y/O NARIZ, PROLONGADA
RESPIRACIÓN PROFUNDA	3	INHALACIÓN Y EXHALACIÓN POR BOCA Y/O NARIZ, PROLONGADA
BOSTEZO	4	ACCIÓN INVOLUNTARIA DE INSPIRAR AIRE LENTA Y PROLONGADAMENTE, ABRIENDO LA BOCA, Y LUEGO EXPIRARLO TAMBIÉN PROLONGADAMENTE Y A VECES CON RUIDO
ESTIRAMIENTO	4	ELONGACION DE BRAZOS, PIERNAS, ESPALDA Y CUELLO
PAUSA	4	PERIODO CORTO DE TIEMPO DONDE NO SE REALIZA ACTIVIDAD
ACOMODO EN SILLA	3	MOVIMIENTOS SUAVES EN LA SILLA QUE NO CAMBIAN POSICIÓN
AJUSTE DE POSICIÓN	4	CAMBIA LA POSICIÓN SENTADO O CONFIGURACIÓN DEL ENTORNO

MOVIMIENTOS BRUSCOS	5	MOVIMIENTOS MÁS RÁPIDOS A LOS NORMALES, QUE CAUSAN MAYOR IMPACTO AUDITIVO DE LOS ELEMENTOS UTILIZADOS EN LA TAREA
PREGUNTAR CUÁNTO FALTA	5	PREGUNTA REALIZADA PARA CONOCER LA HORA, TIEMPO TRANSCURRIDO Y/O REMANENTE DE LA EJECUCIÓN DE LA TAREA

Tabla VIII: Niveles de fatiga observada

El paso final del etiquetado de fatiga es la definición de los niveles de fatiga observados en la muestra experimental. Para esto primero se estandarizan (z-score) los valores obtenidos para cada dimensión de fatiga (KSS, Autoreportada, Rendimiento) y luego se realiza una clusterización con estas 3 dimensiones de fatiga identificadas.

ii) Clusterización de etiquetas

En la definición de los clúster finales de fatiga (niveles), primero se calcula el número óptimo de clusters con el criterio mediante un criterio de evaluación, como por ejemplo Calinski Harabasz. Luego, se aplica un método de clustering, por ejemplo k-means o Gaussian Mixtures Models, con el número de clusters determinado anteriormente (Xu, 2009; Calinski & Harabasz, 1974). Para determinar cuál cluster corresponde a cada nivel se calcula la sumatoria de cada dimensión del centroide del cluster y se ordenan de menor a mayor, siendo el mínimo el nivel más bajo y el máximo el más alto. Esto es válido porque las 3 escalas usadas, fatiga autorreportada, observada y rendimiento, son directamente proporcionales a la fatiga. De lo anterior se consigue determinar el número de niveles de fatiga y la fatiga para cada ventana de observación. Esta información es finalmente agregada a la tabla de datos obtenida en C.7 y con este resultado se pasa a la etapa de análisis de la tarea.

iii) Clasificación

Si bien la clasificación no es aparte de la metodología formal se realiza un análisis prospectivo debido al potencial que presentan los modelos de clasificación. Estos permitirían, con un conjunto grande de datos, realizar predicciones de fatiga y automatizar procesos como el de etiquetado. Utilizando el dataframe con las etiquetas es posible aplicar modelos de aprendizaje supervisado de machine learning.

El primer paso es realizar una selección de características, paso que estima la relevancia de cada una de ella y elimina las que no aporten al proceso de clasificación, esto se realiza con el algoritmo Boruta. Para la clasificación se utilizan tres modelos: Support Vector Machine (SVM), Multi-Layer Perceptron (MLP) y Random Forest (RF). Además, para evitar el sobreajuste se aplica a cada modelo el método de Repeated Cross-Validation (10-fold, repetido tres veces) lo que evita

que el clasificador se sobreentrene. Los resultados se presentan como porcentajes de Accuracy, Precision, Recall, F1-score, Kappa, que son distintas métricas de desempeño de la clasificación.

D.- ANÁLISIS DE LA TAREA

En esta etapa se busca caracterizar la tarea en función de la fatiga que esta produce sobre los sujetos experimentales mediante el análisis de las señales psicofisiológicas capturadas para cada individuo. Como se muestra en la figura 2, esta caracterización contempla dos tareas principales:

- **El análisis del efecto de la tarea sobre cada individuo:** Descripción del impacto de la tarea sobre las señales psicofisiológicas sobre cada sujeto y la agrupación de sujetos con reacciones similares.
- **Descripción de la tarea para la muestra agregada:** Descripción de los patrones comunes observados en los sujetos durante la realización del experimento y los marcadores psicofisiológicos que explican distintos niveles de fatiga.

De esta manera se busca entender el comportamiento de los sujetos experimentales para luego describir los patrones de fatiga comunes observados en toda la muestra y encontrar variables psicofisiológicas que permitan detectar fatiga.

9.- Efecto de la tarea por individuo

Se inicia el análisis con la descripción individual y clasificación de sujetos experimentales en grupos de comportamiento psicofisiológico. Esto resulta necesario debido a la naturaleza heterogénea de los sujetos experimentales, lo que genera una alta variabilidad en las reacciones de los sujetos ante el experimento. Esta variabilidad puede implicar la necesidad de analizar casos particulares de sujetos cuya fisiología resulte muy distinta o simplemente eliminar datos que potencialmente invaliden los resultados generales del experimento. Esta variabilidad podrían ser explicadas por motivos tales como diferencias de edad, estado físico, hábitos como consumo de estimulantes y alimentación, diferentes patologías cardiovasculares y hormonales, etc.

Para lidiar con la variabilidad entre sujetos, se propone la utilización de métodos de clustering basados en las series de tiempo de sus señales psicofisiológicas. Utilizando estas variables se propone el proceso de análisis y clusterizado de sujetos expuesto en la figura 7. El resultado final de esta etapa es la agrupación de usuarios en función de la evolución de sus señales psicofisiológicas durante la ejecución de la tarea experimental.

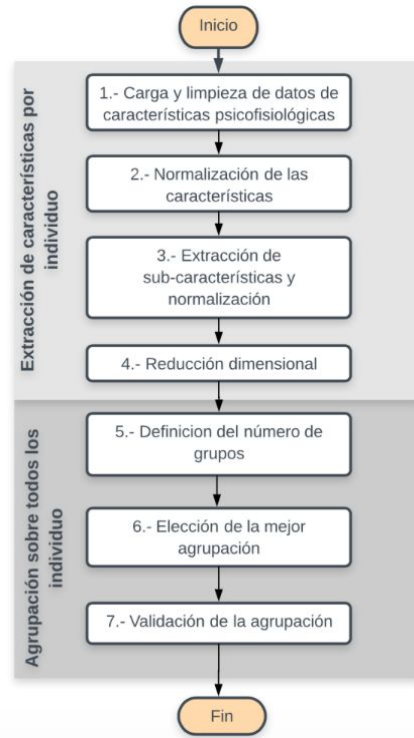


Figura 9: Proceso de clusterización de sujetos según sus señales psicofisiológicas

La figura 9 presenta el proceso de clusterización compuesto por las siguientes etapas: i) Extracción de características por individuos y ii) Agrupación sobre todos los individuos. Estas etapas se explican a continuación:

i) Extracción de características por individuos: Esta etapa comienza con la carga de las series de tiempo de las características psicofisiológicas y un proceso de limpieza para eliminar los registros anómalos que pueden distorsionar las relaciones entre los sujetos (Figura 8.1). Esto se realiza descartando todos los valores de la serie de tiempo que estén fuera del rango definido por el criterio de “inner fence” para cada característica de cada sujeto experimental. Luego se estandariza (z-score) cada subconjunto de datos para cada sujeto (Figura 8.2), llevando todos los datos a un rango de (-1,1).

El proceso continúa con la extracción de subcaracterísticas (Figura 8.3) para lo que se divide la serie de tiempo de cada característica por sujeto en 10 tramos. Para cada uno de estos tramos se extraen 7 estadísticos que caracterizan el tramo (media, mediana, primer cuartil, tercer cuartil, máximo, mínimo y la desviación estándar). De esta forma se obtiene un vector de subcaracterísticas de largo 1470² para cada sujeto. Seguidamente se aplica la estandarización z-score de cada sub característica a través de los sujetos.

² 1470 corresponde a 10 tramos x 7 estadísticos por características x 21 características

Esta etapa finaliza (Figura 8.4) con una reducción dimensional realizada sobre las 1470 subcaracterísticas realizada con el método de análisis de componentes principales (PCA). Esta reducción permite representar las 1470 subcaracterísticas en un hiperespacio de 9 a 14 componentes principales dependiendo de la tarea laboral y la cantidad de sujetos experimentales que se considere. La definición del número de componentes principales es un proceso iterativo en donde se agregan componentes hasta que se logre un 100% de explicación de la varianza de los datos.

ii) Agrupación sobre todos los individuos: Una vez definidas las características por usuario se continúa con la agrupación de usuarios. Para esto se utiliza el método de Gaussian Mixtures Models (GMM), el que utiliza diferentes matrices de covarianzas (kernel) para realizar agrupaciones (diagonal, esférica, tied y full). Como métrica de evaluación de las distintas agrupaciones obtenidas con el GMM se utiliza el *Bayesian Information Criterion* (BIC).

Para definir el *Kernel* a utilizar se realiza una clusterización para cada tipo de kernel y para cada cantidad de clusters, se calcula el BIC de estas configuraciones y se escoge la que minimice este valor. Además, para confirmar la cantidad de clusters se calcula el silhouette value, donde la configuración ganadora es la de mayor puntaje.

Dado que el modelo GMM no es determinístico entrega diferentes resultados por cada ejecución con la misma configuración de parámetros. Para escoger la mejor ejecución, se realizan 10 iteraciones de esta y se escoge la que tiene menor valor de BIC.

Finalmente, esta clusterización es validada realizando una prueba estadística donde la hipótesis nula es que existe homocedasticidad (igualdad de varianzas) entre cada uno de los 10 tramos de cada señal psicofisiológica (HR, HRV, EDA, SCL, SCR, SKT) y de cada cluster con respecto al tramo equivalente agrupando todos los clusters, para esto se usa el test de Levene y además se realiza la prueba de Kruskal Wallis para testear la hipótesis nula de que existe diferencia significativa entre las medias de las señales de los clusters y para cada tramo.

10.- Descripción de la tarea para la muestra agregada

La segunda etapa del análisis busca encontrar patrones comunes de fatiga sobre el total de usuarios y encontrar marcadores psicofisiológicos que expliquen la fatiga de los sujetos experimentales. Para este análisis se realiza mediante una serie de herramientas de análisis estadístico para comprobar 3 hipótesis sobre la relación entre la fatiga y la psicofisiología de los sujetos experimentales. Estas hipótesis se describen a continuación en la tabla IX y se explican a continuación.

Hipótesis	Técnica
1) Existe dependencia estadística entre el nivel de fatiga y los grupos de sujetos experimentales.	Tabla de contingencia, prueba de chi-cuadrado y visualización de datos.
2) Existe una tendencia creciente de la fatiga con el avance temporal de la tarea laboral.	Ajuste lineal y visualización de datos.
3) Existe diferencia significativa en las medianas y varianzas de las características psicofisiológicas entre los niveles de fatiga.	Prueba de Kruskal Wallis, prueba de Levene.

Tabla IX: Hipótesis para la caracterización de la tarea.

1. Existe dependencia estadística entre el nivel de fatiga y los grupos de sujetos experimentales:

Debido a que en esta hipótesis se desea probar la existencia de dependencia entre dos variables categóricas, tales como el nivel de fatiga y la etiqueta de pertenencia a un grupo de cada sujeto experimental (resultado del clustering por comportamiento fisiológico), es que se debe construir una tabla de contingencia que contiene la cantidad de ventanas de tiempo (observaciones) de cada una de las categorías.

Luego se aplica una prueba de chi-cuadrado para determinar si el valor observado de una variable depende del valor observado de otra variable, vale decir, dependencia estadística entre los niveles de fatiga y los grupos de sujetos, finalmente se grafica las distribuciones de fatiga para cada grupo con el objetivo de visualizar estas diferencias y sacar conclusiones. A continuación, se presentan los resultados.

2. Existe una tendencia creciente de la fatiga con el avance temporal de la tarea laboral:

En este análisis se divide el largo temporal de la tarea en 10 tramos y para cada uno se calcula la fatiga promedio, con esto se construye una curva y se realiza un ajuste lineal que representa la tendencia que sigue la fatiga a través del tiempo. Utilizando una visualización de estos datos junto con los valores de los parámetros del ajuste lineal se sacan conclusiones.

3. Existe diferencia significativa en las medianas y varianzas de las características psicofisiológicas entre los niveles de fatiga:

Para determinar significancia estadística en la diferencia de medianas y varianzas de las características psicofisiológicas entre los 3 niveles de fatiga identificados (sin fatiga, baja fatiga y alta fatiga) se realizan pruebas estadísticas conocidas, para este caso se utilizarán las pruebas de Kruskal Wallis para probar diferencia significativa entre medianas y la prueba de Levene para la probar la diferencia significativa

entre las varianzas. Junto con lo anterior se calcula los valores de las medianas y desviación estándar de cada caso y se identifica si existe algún patrón de aumento o disminución entre los valores de las características y el aumento de fatiga laboral.

11.- Resultados y Conclusiones

En esta etapa se debe concluir en al menos 3 puntos: i) El comportamiento de los sujetos ante la tarea, ii) La existencia de niveles de fatiga y iii) los marcadores psicofisiológicos existentes que expliquen la fatiga.

- I. **El comportamiento de los sujetos ante la tarea:** Se busca responder si la tarea ejerce efectos de fatiga distintos en los sujetos (grupos de sujetos) y cuál es el patrón de esta fatiga en los grupos. En otras palabras si existen patrones de fatiga distintivos según grupos de usuarios. Esto se puede responder con la hipótesis 1.
- II. **La existencia de niveles de fatiga y su tendencia en el tiempo:** Se busca responder si los niveles encontrados tienen validez y significancia estadística, además del comportamiento en el tiempo de estos niveles. Esto se puede responder con la hipótesis 2 y 3.
- III. **Los marcadores psicofisiológicos existentes que expliquen la fatiga:** Se busca concluir cómo la tarea afecta las señales psicofisiológicas de los usuarios y si existen patrones agregados (por grupos) que se relacionen con la aparición de fatiga. Esto se puede responder con las hipótesis 3.

3.- APLICACIÓN DE LA METODOLOGÍA: CONDUCCIÓN

En esta capítulo se presenta la metodología presentada en la sección 1 aplicada a una tarea de conducción profesional en terreno, específicamente la ruta Santiago-Los Bronces, correspondiente al traslado de trabajadores de la minería. Esta presentación se hará siguiendo la agrupación por etapas introducida en el capítulo anterior. Para cada etapa se desarrollarán las actividades correspondientes con datos capturados en el experimento de conducción.

A.- DEFINICIÓN DE LA TAREA

1.- Caracterización de la tarea

A continuación se define la tarea según lo definido en la sección 1.A., mediante una descripción, sus actividades y métricas de rendimiento.

Descripción: Esta labor consiste en la conducción profesional de buses de traslado de pasajeros, ida y vuelta, entre Santiago y la Mina Los Bronces. Esta labor es no monótona, no repetitiva y continua, Tabla X. De acuerdo a la literatura estudiada, esta tarea es de atención sostenida (Michael R, 2006). La atención sostenida es la necesidad de mantener el foco en un estímulo continuo. En el presente caso, la tarea requiere la capacidad de mantener el foco en la conducción ante estímulos continuos y variados desde distintas fuentes, tales como: ambientales, que requieren atención visual del camino, ajustando los movimientos físicos del conductor para operar el vehículo; y administrativas, que requieren reportes de estado del camino.

Actividades que componen la tarea: Estas actividades se inducen de las visitas a terreno y de la información entregada por la misma empresa.

- Inicio de la conducción.
- Fin de la conducción.
- Detención por señalética.
- Conducción en Curvas.
- Conducción en Ascenso.
- Conducción en Descenso.
- Conducción en Ciudad.
- Conducción en Camino de montaña.
- Control de la velocidad del bus.
- Conducción en zonas de riesgo.

Tarea/Hito	Actividades	Monotonía	Repetitividad	Continuidad
Conducción: Conducir bus con pasajeros, entre Santiago y Mina Los Bronces.	Todo lo que involucra una conducción profesional en camino de ciudad y montaña	No Monotona	No Repetitiva	Continua

Tabla X: Caracterización de la Actividad de Conducción

Métricas de rendimiento: Se definieron métricas en relación con la aceleración lateral y longitudinal de la cabina, ya que se consideraron que están en directa relación con un estado de fatiga durante la tarea de una conducción segura y confortable para los pasajeros, Tabla XI. La magnitud resultante de las aceleraciones lateral y longitudinal se normaliza entre 0 y 1 con respecto a toda la serie de tiempo de aceleraciones de cada sujeto y subactividad.

Actividad	Métrica de efectividad	Métrica de eficiencia
A) Conducción	Aceleraciones normalizadas sobre 0,5	Cantidad de veces que la aceleración obtuvo valor superior a 0,5, durante la ventana de tiempo.

Tabla XI: Métricas de rendimiento para la actividad de conducción.

A partir de las métricas de efectividad (Ne) y eficiencia (Fe), se define un indicador de errores (Ie), de la siguiente forma $Ie = p \cdot Fe + (1 - p) \cdot Ne$, donde p es un ponderador que da cuenta de la importancia que se asigna a la frecuencia de aceleraciones sobre 0,5. En este caso se estableció $p=0,7$, en base a la estimación de juicio experto, respecto a importancia asignada a la frecuencia de errores por sobre la cantidad de errores, ya que la ocurrencia reiterada de errores se acerca más a la definición del fenómeno de fatiga (Ma,2018; Fu,2016; Lin, 2013; He, 2013; Dahir, 2018).

Para la actividad de conducción, la métrica de efectividad se definió como la suma de las diferencias (Ne) entre los valores de aceleración y 0,5, de cada peak sobre las 0,5. La métrica de eficiencia fue definida como la frecuencia de estos eventos (Fe). Estas métricas buscan cuantificar cuán brusca fue la conducción con respecto al estilo de conducción propio y de cómo conducen los demás sujetos en el mismo lugar de la ruta, medida como una magnitud de las aceleraciones normalizadas sobre 0,5, y establecer la frecuencia con que se repitió esta situación. Por ejemplo, en la figura 10, $Ne = a + b$ y $Fe=2$

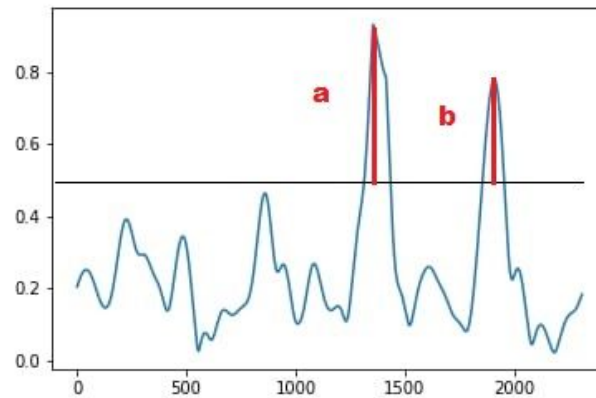


Figura 10: Figura explicativa de cálculo de métrica de rendimiento a partir de la aceleración durante la ruta.

Por otro lado, la etiqueta de rendimiento (E_r) para el modelo se definió como una combinación lineal del indicador de errores (I_e), tiempo medio de ejecución (T_{em}) y varianza del tiempo de ejecución ($Var(T_e)$), mediante esta relación: $E_r = \alpha \cdot I_e + \beta \cdot T_{em} + \gamma \cdot Var(T_e)$. La conducción, al ser una tarea no repetitiva, implica que $\beta = \gamma = 0$ y $\alpha = 1$; por lo tanto, la etiqueta de rendimiento corresponderá al indicador de errores definido para conducción.

2.- Caracterización del entorno de ejecución de la tarea

A continuación se caracterizan los entornos de ejecución de las tareas descritas en el punto anterior. De acuerdo a la metodología la descripción se realizará en términos del entorno directo e indirecto de ejecución.

Entorno directo: La cabina del conductor define los estímulos directos a los que se encuentra sometido. Define también las condiciones de confort físico bajo las cuales se llevan a cabo las tareas de conducción. Las figuras 10.a y 10.b muestran la cabina con sus dimensiones. Si bien el tamaño de la cabina cambia por modelo de bus, las fotos dan cuenta del espacio real en las que el chofer conduce. La figura 11 muestra el tablero de conducción del bus, en este tablero se encuentran los indicadores mecánicos de funcionamiento del bus, velocidad y control de las puertas.

Las condiciones de entorno afectan de manera importante tanto el rendimiento como la fatiga sufrida por el trabajador. Dentro de las variables relevantes descritas en trabajos relacionados para el caso particular de la conducción profesional se cuentan: vibración del

bus, concentración de CO2 en la cabina, temperatura ambiental en la cabina, nivel de ruido en la cabina, luminosidad en la cabina. Se muestra en la figura 11.b un espacio inmediatamente detrás de la cabecera del asiento del chofer.

a)



b)



Figura 11: a) Cabina de conducción. b) Compartimento para los sensores de contexto.

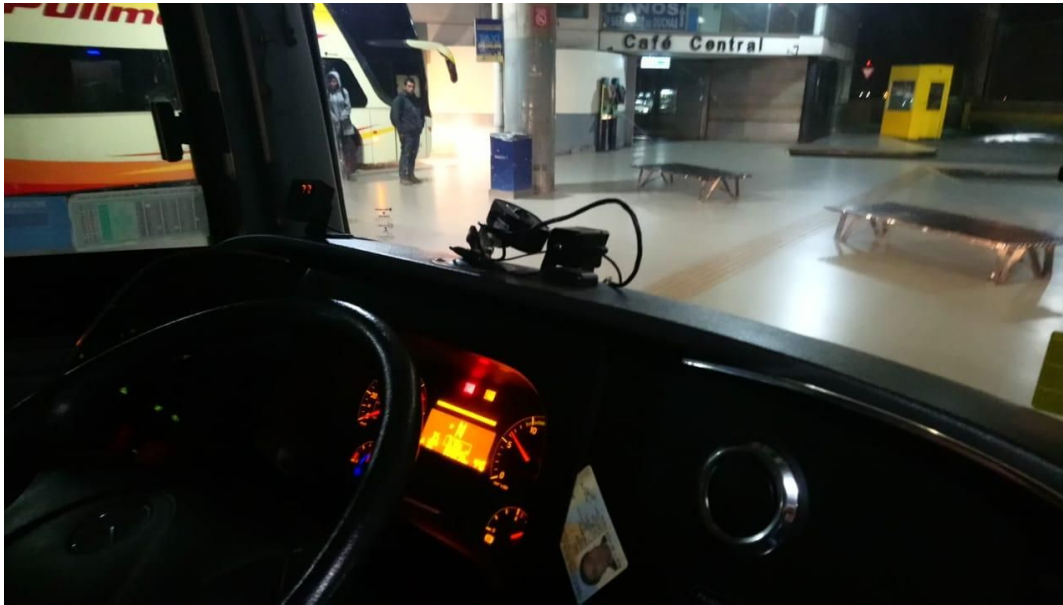


Figura 12: Tablero de conducción del bus.

Entorno indirecto: Como entorno indirecto se considera la ruta de conducción, es decir, el camino por donde se maneja el bus y su entorno. La ruta se ha descompuesto en dos tramos: i) tramo específico del conductor "Terminal - Las puertas" que corresponde al tramo en ciudad y ii) Tramo común "Las puertas - Los Bronces" que es el tramo en montaña.

El primer tramo no es común a todos los conductores y depende del punto de partida del bus y de los paraderos. En efecto, los choferes comienzan su jornada al llegar al lugar de estacionamiento del bus. En ese punto realizan una inspección de la máquina (la "vuelta del perro") y luego parten su conducción hasta el paradero donde se recoge a los trabajadores. Los lugares de estacionamiento son: el taller (ubicado en calle San Borja) y el llamado "Patio Ferro", ubicado en la intersección de Ramón Subercaseaux con San Vicente. Desde el lugar de estacionamiento cada bus se dirige al paradero correspondiente a su ruta. Tablas XII y XIII listan los paraderos de tramos de ciudad.

SERVICIO TURNO H SANTIAGO LOS BRONCES					
Servicio	Tipo Servicio	N° INT	Año	PU	Vehiculo
N° 1 Puente Alto	Turno H AA	3251	2015	GLYZ-57	Bus 14 mts
N° 2 San Bernardo	Turno H AA	3239	2015	HFVR-31	Bus 14 mts
N° 3 San Miguel	Turno H AA	3244	2015	HFVR-36	Bus 14 mts
N° 4 Maipú	Turno H AA	3264	2015	GPGF-88	Bus 14 mts
N° 5 Pudahuel	Turno H AA	Termina	2015	#N/A	Bus 14 mts
N° 6 Quilicura	Turno H AA	3212	2015	GKPD-90	Bus 14 mts
N° 7 La Pintana	Turno H AA	3185	2015	GKPD-86	Bus 14 mts
N° 9 Terminal	Turno H AA	3261	2015	GPGF-84	Bus 14 mts
N° 10 Las Puertas	Turno H AA	3216	2015	GKPD-95	Bus 14 mts
N° 11 Sta Rosa-Pío Nono	Turno H AA	3217	2015	GKPD-96	Bus 14 mts
N° 12 Superv.Placa La Florida	Turno H AA	3218	2015	GKPD-97	Bus 14 mts
N° 13 Pío Nono-Las Puertas	Turno H AA	3211	2015	GKPD-89	Bus 14 mts
N° 14 Supervisores Vespucio Colon	Turno H AA	3213	2015	GKPD-91	Bus 14 mts

Tabla XII: Descripción del punto de partida y las máquinas que operan en el turno H

SERVICIO MONO OPERADOR SANTIAGO LOS BRONCES					
Servicio	Tipo Servicio	N° INT	Año	PU	Vehiculo
Contratistas (Terminal - Bronces)	Turno A/B	3242	2015	HFVR-34	Bus 14 mts
Terminal 1	Turno H MOP	3412	2016	HYXB-67	Bus 2P
Terminal 2	Turno H MOP	3413	2016	HYXB-68	Bus 2P
Terminal 3	Turno H MOP	3414	2016	HYXB-69	Bus 2P
Terminal 4	Turno H MOP	3416	2016	HYXB-71	Bus 2P
Terminal 5	Turno H MOP	3417	2016	HYXB-72	Bus 2P
Terminal 6	Turno H MOP	3418	2016	HYXB-73	Bus 2P
La Florida-14 Vicuña Mackena	Turno H MOP	3411	2016	HYXB-66	Bus 2P
Maipu	Turno H MOP	3415	2016	HYXB-70	Bus 2P

Tabla XIII: Descripción del punto de partida y las máquinas que operan en el turno mono operador (Terminal es el Terminal de buses San Borja).

El segundo tramo es común para todos los buses, correspondiente a “Las puertas - Los Bronces”. Este tramo se muestra en la Figura 13, corresponde al segmento de montaña y termina en el campamento “Pérez Caldera” o en el hotel “Placa”.

A lo largo de esta ruta de montaña existe una serie de situaciones que presentan potenciales riesgos para el conductor. Estas situaciones de riesgo se recogen y documentan sistemáticamente por la empresa de transporte de pasajeros en el reporte IDPIER (Identificación de Potenciales incidentes en la Ruta). La Figura 14 muestra el ejemplo de documentación de una de estas situaciones de riesgo. En base a la información entregada por la empresa es posible etiquetar el camino con esta información, lo que permitiría correlacionar y comparar estas situaciones de riesgo con la fatiga de los conductores.

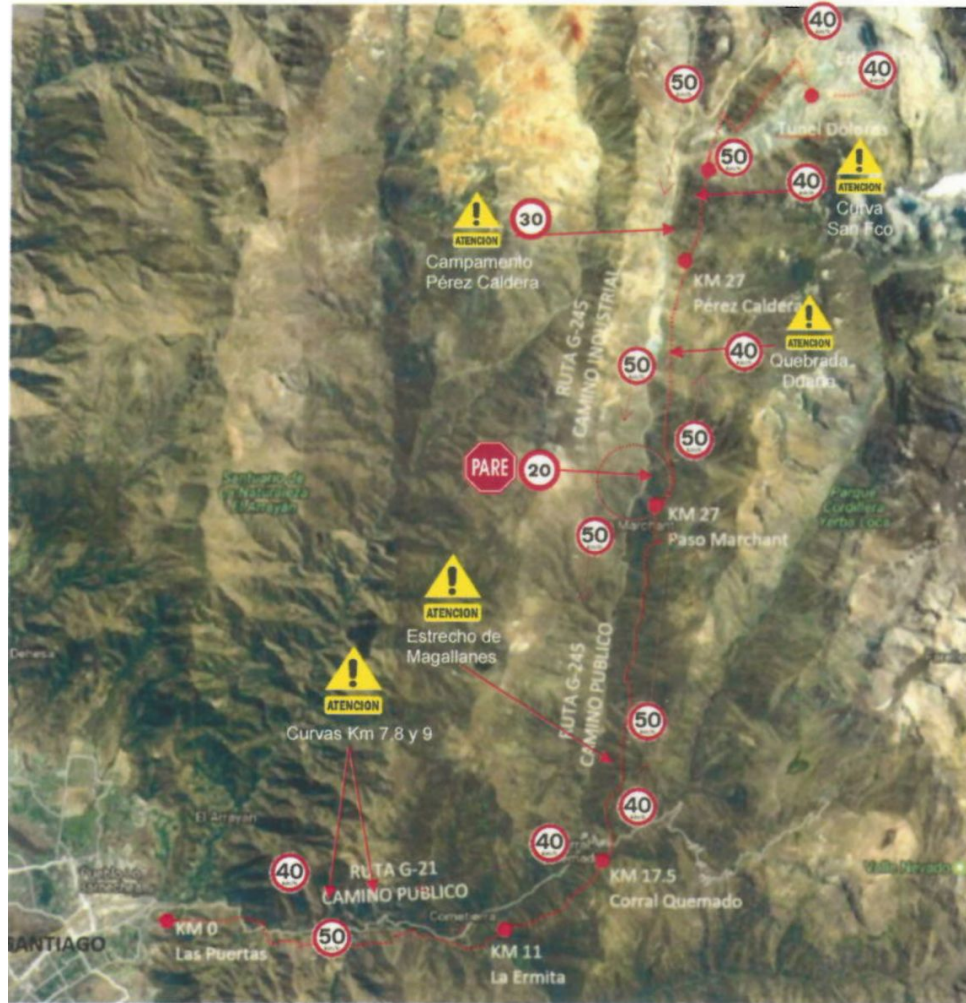


Figura 13: Vista satelital de la ruta "Las puertas - Los Bronces", camino de montaña.

<p>G-245 KM-19,5 Sector motos</p>	<p>las</p> <p>Socavón al costado del camino, peligro de caída de vehículos que transitan por el lugar</p>	<p>Mejorar condición de camino</p>	
--	---	------------------------------------	--

Figura 14: Documentación de una situación de riesgo en el tramo "Las puertas - Los Bronces".

3.- Definición de la muestra

De acuerdo a lo explicado en el punto 1.A.3, a continuación se define la muestra en función del tamaño muestral y el número de observaciones: i) tamaño muestral: 30 usuarios (26 con turno 5x2, tres con turno 7x7 y uno con turno 6x3; ii) número de observaciones: 43.017 en total para las señales psicofisiológicas.

	PROMEDIO	DESV EST			
EDAD AÑOS	47,5	21,3			
	MASCULINO	FEMENINO			
GENERO	30 / 100%	0 / 0%			
	SOLTERO	CASADO	DIVORCIADO		
ESTADO CIVIL	11	18	1		
	SIN	CON			
HIJOS	0 / 0%	30 / 100%			
	BASICA	MEDIA	SUPERIOR	POSTGRADO	
EDUCACION	1	26	3	0	
	SIN TRABAJO	A PLAZO FIJO	POR PROYECTO	INDEFINIDO	INDEPENDIENTE
TIPO DE CONTRATO	0	0	0	30	0
	SIN TRABAJO	MENOS 1 AÑO	1-5 AÑOS	5 AÑOS O MAS	
PERIODO TRABAJO EN EMPRESA	0	0	9	21	
	SIN EXPERIENCIA	MENOS 1 AÑO	1-5 AÑOS	5 AÑOS O MAS	
EXPERIENCIA LABORAL	0	0	0	30	
	SI	NO			
ACTIVIDAD FISICA SEMANAL ³	19 / 61%	12 / 29%			
	SI	NO			
¿DURMIO BIEN ANOCHE? ⁴	27 / 87%	4 / 13%			

³ Actividad física semanal: Se refiere a practicar algún deporte o ejercitar en gimnasio, al menos una vez por semana.

⁴ ¿Durmió bien anoche?: Se pregunta directamente para establecer la percepción de un buen descanso durante el sueño.

	PROMEDIO	DESV EST			
CANTIDAD HORAS DE SUEÑO ⁵	4,05	1,09			

Tabla XIV: Descripción de la muestra y observaciones para conducción.

4.- Preparación de artefactos y materiales

De acuerdo a lo expuesto en el punto 1.A.4, los artefactos y materiales utilizados son:

- I. **Medidas fisiológicas:** Biomonitor.
- II. **Medidas Contextuales**
 - A. **Contexto de entorno directo:** sensor de contexto (explicado en el punto 1.A.4)
 - B. **Contexto de entorno indirecto:** GPS, subtareas, cuestionarios y registro de respuestas en ruta.
- III. **Formularios legales:** Consentimiento informado (explicado en el punto 1.A.4)

B.- DEFINICIÓN DEL EXPERIMENTO

De acuerdo con la metodología en esta sección se diseña e implementa el experimento. Concretamente se define el protocolo experimental, se ejecutan los experimentos y se capturan los datos experimentales.

5.- Definición del protocolo experimental

A continuación se describe el protocolo experimental siguiendo los lineamientos expuestos en el punto B.5

Antes del experimento

- 1.- Experimentador debe presentarse al lugar de trabajo con baterías cargadas y todos los dispositivos y aplicaciones dispuestas para una actividad experimental fluida.
- 2.- Bienvenida, descripción general del experimento y completar consentimiento informado, cuestionarios y cualquier otro documento, si corresponde. Según los tiempo disponibles, el experimentador puede dejar el llenado de información para otro momento durante la sesión experimental.
- 3.- Instalar sensores al participante. Verificar la comodidad y funcionamiento.

⁵ Cantidad horas de sueño: Corresponde a las horas de sueño de la noche, antes del inicio de la jornada laboral, no incluye las horas de sueño del descanso antes de iniciar el tramo de vuelta.

- 4.- Instalar el sensor de contexto directo en la cabina del conductor. Validar su funcionamiento correcto.
- 5.- Considerar período de adaptación a sensores y etapa de nivelación, si aplica.
- 6.- Instruir al participante sobre la aplicación del subareas de capacidad de atención y autorreporte de fatiga, N-back y KSS, respectivamente.
- 7.- Registro de línea base. Etapa de medición de 5 minutos sin estímulos con el participante sentado en el área de trabajo, es decir, sentado frente al volante.

Durante el experimento

- 1.- Dar instrucciones para la ejecución de la tarea.
- 2.- Recordar que se realizarán N-Back y KSS, durante la tarea.
- 3.- Inicio la ejecución de la tarea del experimento.
- 4.- Supervisar ejecución de la tarea y funcionamiento de los sensores, durante todo el experimento.
- 5.- Registrar tiempo e información relevante de eventos.
- 6.- Registrar gestos asociados a fatiga del participante.
- 7.- Una vez iniciada la tarea, el participante debe responder el cuestionario n-back y KSS, cada 15 minutos.
- 8.- Participante termina la tarea al completar una tramo de la ruta, ya sea ida o vuelta.

Después del experimento

- 1.- Toma de línea base. Etapa de medición de 5 minutos sin estímulos con el participante sentado en el área de trabajo.
- 2.- Indicar término de la actividad.
- 3.- Detener y retirar los sensores. Se finaliza la toma de datos.
- 4.- Preguntar y reportar cualquier situación relevante para el análisis posterior (situación anómala, incómoda, estrés, sueño, etc.).
- 5.- Despedida del participante y agradecimiento por participación.
- 6.- Descarga y almacenamiento de archivos de las distintas fuentes de datos.

6.- Ejecución del protocolo experimental

La captura de datos se realizó según lo explicado en el punto B.5, con una duración completa por experimento de aproximadamente 3 horas. La fecha de inicio fue el 12/12/2018 y la fecha de término el 28/02/2019 teniendo un total de 38 días de medición con 75 muestras (37 capturas de subida y 38 de bajada). Salvo por dos casos excepcionales, los viajes de ida se iniciaron

todos entre las 4:12 y las 6:06 horas, y terminaron entre las 7:12 y las 8:26 horas. Por su parte, los viajes de vuelta se iniciaron entre las 14:08 y las 17:07 horas, y terminaron entre las 17:05 y las 21:25 horas.

Una de las dificultades que se presentaron fue que debido a las condiciones de terreno (mucho polvillo en el aire) y a lo extenso de las mediciones, el computador usado para las mediciones se calentaba mucho, por lo que consumía más rápido la batería y se apagaba antes de concluir la ruta. Esto se solucionó adquiriendo un computador ThinkPAD de Lenovo, adaptado para estas condiciones. Por este mismo motivo fue necesario repetir la medición en siete conductores.

Otro aspecto que es importante mencionar es que para los conductores es muy difícil actuar como si no los estuvieran midiendo. Aún cuando se les dio la instrucción explícita de no interactuar con el experimentador, conversaban con él, lo que afecta las medidas observadas que se van registrando. Además, entre ellos bromeaban con el asunto de ser medidos, por lo que posiblemente se encuentre un sesgo de respuesta en las medidas autoreportadas. Fuera de estos puntos los experimentos se realizaron sin mayores dificultades. En las figuras 15 y 16 se observa la medición realizada a los conductores con la pulsera BioMonitor y el sensor de contexto adosado a la cabina.



Figura 15: Conductores de la empresa de transporte de pasajeros siendo medidos en la ruta “Las puertas - Los Bronces”, camino de montaña.



Figura 16: Sensor de contexto instalado en la cabina justo detrás del asiento del conductor.

C.- PREPARACIÓN DE DATOS

Antes de utilizar los datos capturados para la caracterización de la tarea y la descripción de la fatiga, es necesario examinarlos y prepararlos para la aplicación de técnicas estadísticas y de aprendizaje de máquinas. A continuación se presentan los resultados obtenidos para los procesos de “Limpieza y procesamiento de datos” y “Etiquetado de Fatiga” para la actividad de conducción.

7.- Limpieza y procesamiento de datos

Para esta actividad se aplica la etapa C.7 de la metodología. Debido a la complejidad de captura y procesamiento de los datos propios de una medición en terreno, donde el ruido de las señales, principalmente por el movimiento, es alto, existe una pérdida de datos. De esta manera, de los 30 conductores medidos, tanto en el viaje de subida como en el de bajada, se logra procesar correctamente 24 capturas para la subida y 20 capturas para la bajada, teniendo, por lo tanto, una pérdida de datos del orden de 20-35%.

El algoritmo utilizado para procesar PPG y Temperatura presentaron un buen rendimiento. Sin embargo, en los casos en que existe mucho ruido en la señal, el algoritmo de procesamiento de EDA muestra dificultades como, por ejemplo, la imposibilidad de detectar y eliminar valores fuera de

rango que interfieren en el cálculo de las variables secundarias del EDA (componente tónica y fásica) por lo que fue necesario eliminar estos valores fuera de rango a mano y caso a caso.

Para la extracción de características para esta tarea se considera un largo de ventana de cinco minutos con un deslice de diez segundos como propone la metodología. Éstas se calculan para cada uno de los sujetos, obteniéndose un total de 43.017 observaciones para la tarea laboral de conducción.

8.- Etiquetado de fatiga

i) Etiquetado de fatiga para las ventanas de observación: Para las ventanas de cinco minutos de largo con un deslice de diez segundos, se obtienen las siguientes etiquetas (Ver etapa ocho de la metodología):

- **Fatiga autoreportada:**

Para las respuestas del conjunto de participantes de la tarea laboral de conducción se calcula el estadístico Alfa de Cronbach. Este da un valor de 0,61, por lo que no se valida la consistencia interna de las respuestas del cuestionario KSS, lo que indica que los conductores no respondieron adecuadamente su real percepción de fatiga; por lo tanto, este resultado no se ocupa para el etiquetado final de fatiga. La Tabla XV presenta el resumen de los resultados del KSS.

Etapa	Media	Intervalo de confianza	
1	2,25	0,98	3,52
2	2,31	0,96	3,66
3	2,22	0,87	3,57
4	2,22	0,93	3,51
5	2,22	0,98	3,46
6	2,31	0,96	3,66
7	2,39	0,97	3,81
8	2,44	1,14	3,74
9	2,47	1,19	3,75
10	2,47	1,19	3,75

11	2,53	1,25	3,81
12	2,47	1,19	3,75
13	2,44	1,14	3,74
14	2,47	1,17	3,77
Alfa de Cronbach		0,61	

Tabla XV: Resumen de los resultados del KSS para la tarea de conducción.

- **Fatiga observada:** Estudio de los gestos del participante durante la prueba por parte del experimentador. Dado las condiciones de terreno, no pudo implementarse, ya que la ubicación del experimentador y las posiciones que adoptan los conductores durante la tarea, no permitían evaluar adecuadamente los gestos. Por lo tanto, este resultado no se ocupa en el etiquetado de fatiga.
- **Rendimiento:** Etiqueta de rendimiento (E_r) cuenta con 3 dimensiones, 1 basada en los errores y 2 en el tiempo de ejecución.
 - a. I_e : Indicador de error.
 - b. T_{em} : Tiempo medio de tarea.
 - c. $Var(T_e)$: Varianza del tiempo de ejecución

Para calcular estas dimensiones, el cálculo comienza con el preprocesamiento de las señales del sensor de aceleración longitudinal y lateral usando las técnicas de filtro de pasa altos (orden = 1 y cutoff = 0.003) y pasa bajos (orden = 5 y cutoff = 3.6), luego una eliminación de outliers con outer-fence (alfa = 4.5) y finalmente 12 iteraciones del filtro de mediana móvil (orden = 5) y filtro de promedio móvil (orden = 39). Luego, con ambas aceleraciones, se calcula la magnitud de la aceleración resultante y se normaliza entre 0 y 1 con MinMax con respecto a la señal de cada sujeto y subactividad de la tarea, para eliminar la diferencia entre estilos de conducción entre los sujetos, finalmente se calcula las métricas de N_e y F_e para cada ventana de tiempo (de 5 minutos de largo y 10 segundos de avance) como se ejemplifica en la sección A.1 y la Figura 10.

$E_r = \alpha \cdot I_e + \beta \cdot T_{em} + \gamma \cdot Var(T_e)$, donde $\alpha = 1$, $\beta = \gamma = 0$. Debido a que la conducción es una tarea no repetitiva, no pueden considerarse los factores de tiempo de ejecución, y por lo tanto tampoco su varianza.

El valor I_e se calcula como: $I_e = p \cdot F_e + (1 - p) \cdot N_e$ con $p=0,7$. Luego, este indicador se normaliza entre los valores 0 y 1. A continuación se presenta una tabla resumen de los valores

mínimos, máximos, promedio y desviación estándar de los resultados de etiqueta obtenidos (tabla XVI).

	<i>lr</i>
MINIMO	0
MAXIMO	1
PROMEDIO	0,2728
DESV. EST.	02633

Tabla XVI: Resumen de los resultados de fatiga de rendimiento para la tarea de conducción.

Finalmente, producto de que no se puede validar la etiqueta autorreportada y observada, sólo se utiliza la etiqueta de rendimiento para la clusterización. Sin embargo, en el punto 2.D se explotará y explicará en detalle el uso de una métrica de fatiga psicofisiológica.

ii) Clusterización de etiquetas

Para analizar esta tarea, la jornada laboral se dividió en 4 subactividades en función del contexto de conducción. Es decir, si es el viaje de ida o de vuelta de la mina los bronces y si se está en el segmento de ciudad o montaña, teniendo como punto de división la entrada al camino farellones (Lat: -33.366335, Long: -70.494757). De este modo, resultan las subactividades de *ida ciudad*, *ida montaña*, *vuelta ciudad*, *vuelta montaña*. Esta categorización de subactividades se considera en el análisis desde aquí en adelante.

De acuerdo a lo indicado en secciones anteriores, la clusterización se realiza sólo con la dimensión de etiqueta por rendimiento. En primer lugar, se calcula el número óptimo de clusters para cada subactividad con el criterio Calinski Harabasz, cuyas curvas se muestran en la Figura 17, que para cada uno de los cuatro caso da un valor óptimo de $k = 3$, el punto máximo de la curva. Para validar este número se calcula 100 veces el clustering con el método de Gaussian Mixtures Models. Luego, se agrupa utilizando el mismo método. Para determinar cuál cluster corresponde a cada nivel se considera el valor de los rangos de menor a mayor, siendo el mínimo el nivel más bajo y el máximo el más alto. Esto es válido porque la escala usada, fatiga por rendimiento, es directamente proporcional a la fatiga, ya que ésta es un factor de reducción de la productividad de una tarea y de la mala calidad de sus resultados (Aryal., et al., 2017). Además, en la Figura 18 se presenta el resultado de la clusterización para los cuatro casos. Cada elemento de la nube de

puntos corresponde a una ventana de tiempo de cinco minutos a la que se ha etiquetado con alguno de los tres niveles. De los gráficos se observa que la mayor parte del viaje los conductores no presentan fatiga.

Para observar la tendencia de los niveles etiquetados de fatiga en cada participante se realiza un mapeo utilizando la geolocalización. A continuación, se presentan ejemplos de etiquetas para algunos participantes (Figura 13): 1=sin fatiga; 2=Nivel bajo y 3=Nivel alto en las Figuras 19 a 22.

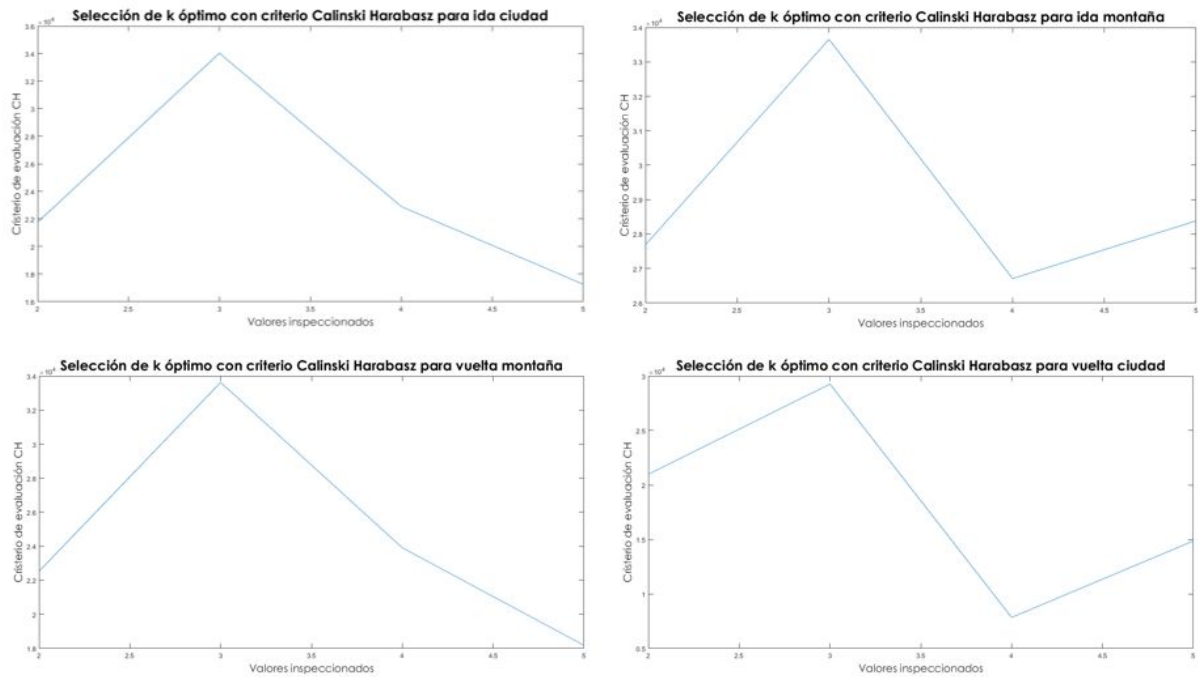


Figura 17: Curvas de criterio Calinski Harabasz para calcular el k óptimo para la las cuatro subactividades de la tarea de conducción.

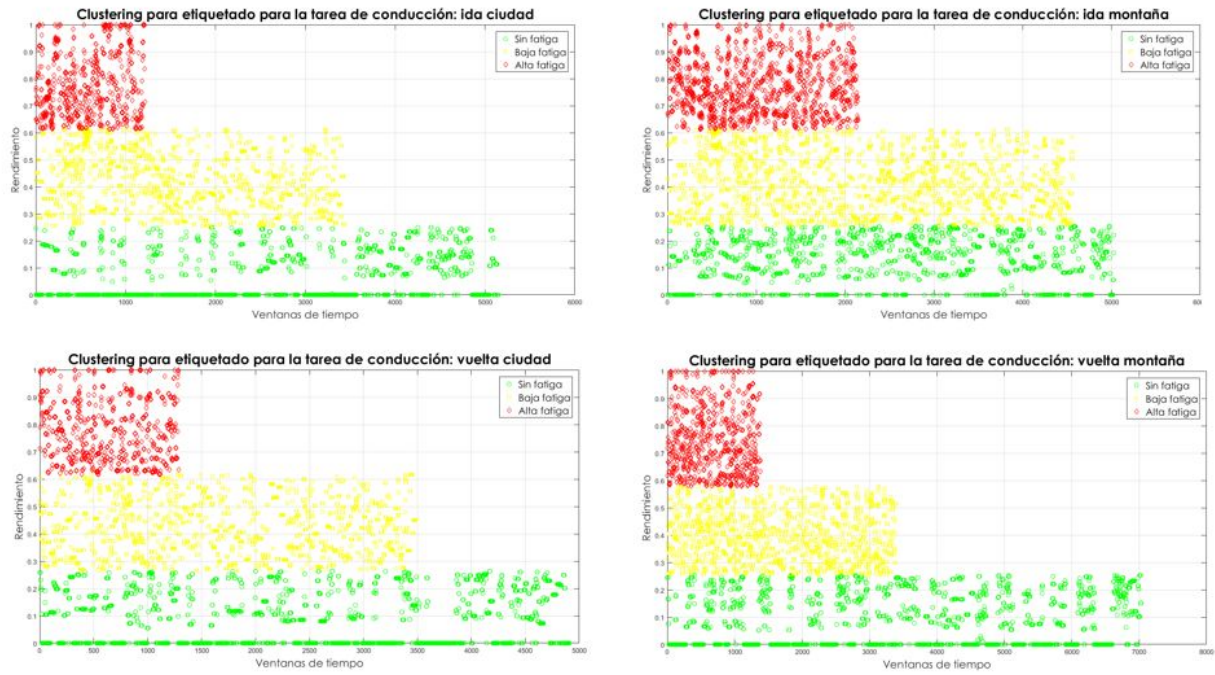


Figura 18: Clustering realizado para las muestras de la tarea de conducción para las cuatro subactividades que permite agrupar las ventanas en 3 niveles distintos de fatiga.

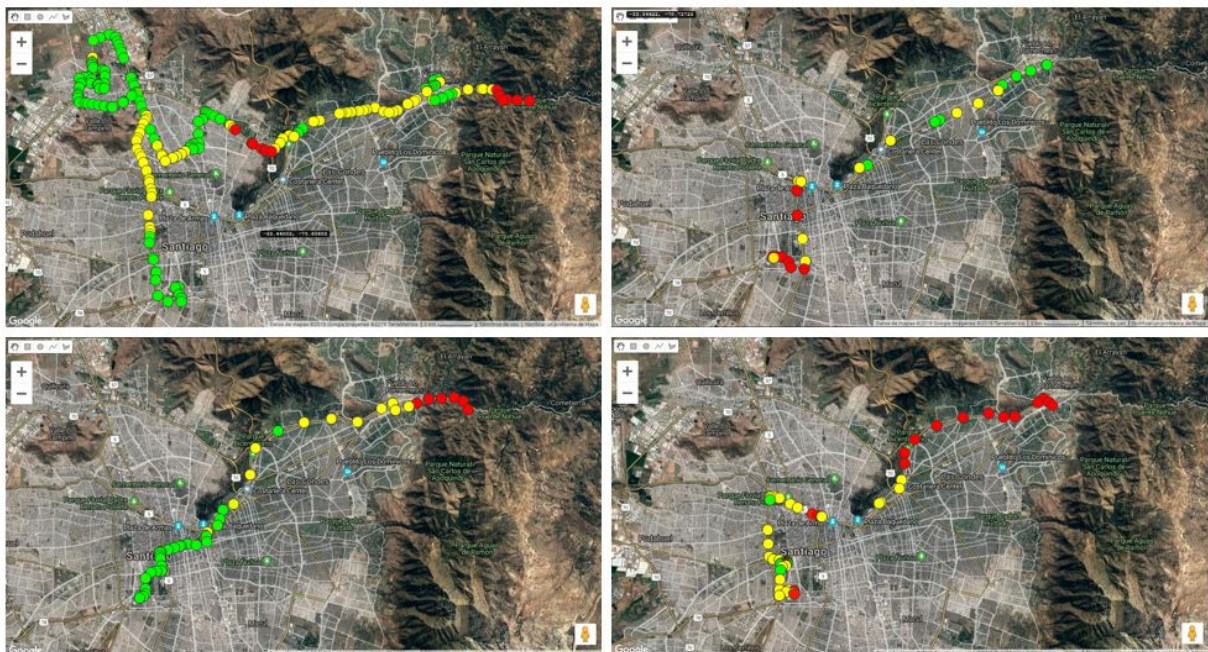


Figura 19: Ejemplo de etiquetado de fatiga por usuarios en tramo ida ciudad. Participantes 28 (arriba izquierda), Participante 17 (arriba derecha), Participante 5 (abajo izquierda), Participante 11 (abajo derecha).

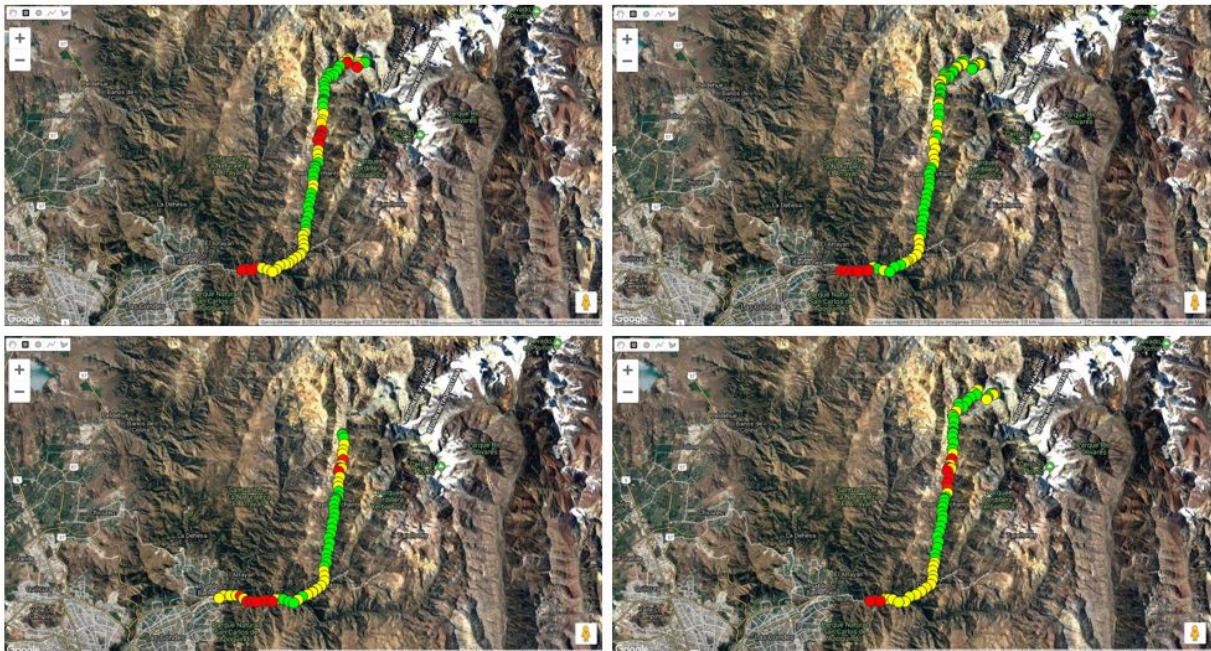


Figura 20: Ejemplo de etiquetado de fatiga por usuarios en tramo ida montaña. Participantes 24 (arriba izquierda), Participante 5 (arriba derecha), Participante 10 (abajo izquierda), Participante 15 (abajo derecha).

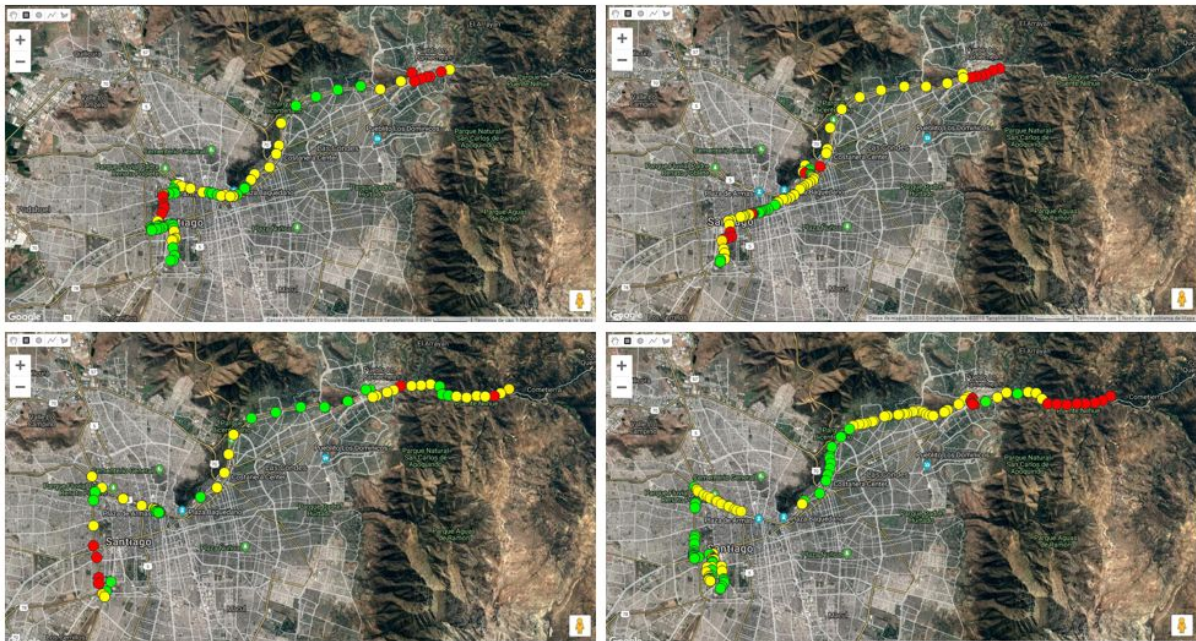


Figura 21: Ejemplo de etiquetado de fatiga por usuarios en tramo vuelta ciudad. Participantes 29 (arriba izquierda), Participante 24 (arriba derecha), Participante 14 (abajo izquierda), Participante 8 (abajo derecha).

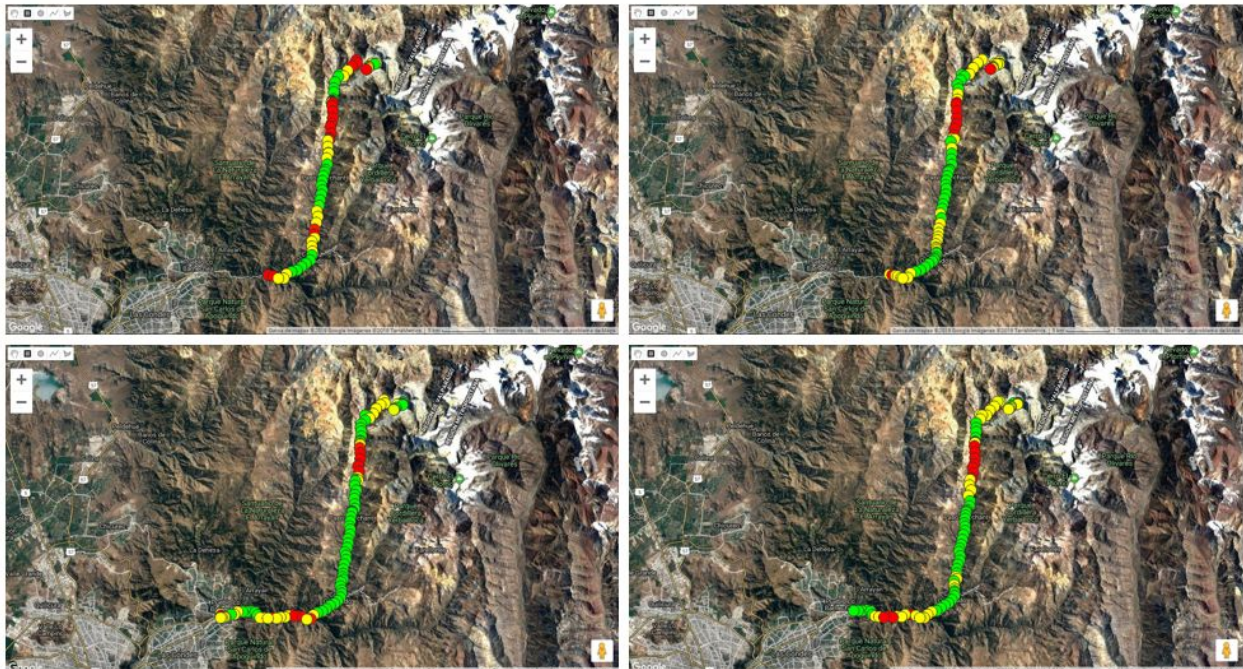


Figura 22: Ejemplo de etiquetado de fatiga por usuarios en tramo vuelta montaña. Participantes 8 (arriba izquierda), Participante 12 (arriba derecha), Participante 21 (abajo izquierda), Participante 28 (abajo derecha).

iii) Clasificación

A cada una de las cuatro subactividades se aplican modelos de clasificación siguiendo los siguientes pasos: Al dataframe de características psicofisiológicas se aplica el algoritmo selector Boruta que arroja en todos los casos que todas las características son relevantes por lo que no se desecha ninguna, en la Anexo I se presenta cada uno de los gráficos de orden de relevancia de las características. Luego, se utilizan dos modelos de clasificación: Support Vector Machine (SVM), y Multi-Layer Perceptron (MLP). Para evitar el sobreajuste se aplica a cada modelo el método de Repeated Cross-Validation (10-fold, repetido tres veces). Luego se agregan a las características psicofisiológicas las medidas tomadas por el sensor de contexto para ver cómo mejora la clasificación incorporando esta información. Se toman como características la media y la desviación estándar de: la temperatura y humedad ambiental, de la concentración de CO2, del ruido ambiental y de la resultante de la aceleración.

En las Tablas XVII a XX se resume el desempeño de cada modelo para cada subactividad para la clasificación de las características de las señales psicofisiológicas considerando las medidas del sensor de contexto como no.

	Accuracy	F1	Kappa
--	----------	----	-------

INFORME VII: INFORME FINAL

	Sin contexto	Con contexto	Sin contexto	Con contexto	Sin contexto	Con contexto
SVM	75,75%	89,93%	68,28%	87,26%	55,46%	82,66%
MLP	63,12%	80,34%	51,69%	76,52%	30,95%	66,61%

Tabla XVII: Resumen de los modelos de clasificación para la tarea de conducción tramo ida ciudad

	Accuracy		F1		Kappa	
	Sin contexto	Con contexto	Sin contexto	Con contexto	Sin contexto	Con contexto
SVM	68,13%	88,71%	65,29%	88,51%	48,28%	82,23%
MLP	57,24%	78,58%	45,39%	75,58%	28,71%	66,14%

Tabla XVIII: Resumen de los modelos de clasificación para la tarea de conducción tramo ida montaña

	Accuracy		F1		Kappa	
	Sin contexto	Con contexto	Sin contexto	Con contexto	Sin contexto	Con contexto
SVM	71,42%	89,77%	63,48%	87,62%	47,72%	82,35%
MLP	60,23%	80,91%	44,04%	77,15%	22,49%	67,57%

Tabla XIX: Resumen de los modelos de clasificación para la tarea de conducción tramo vuelta ciudad

	Accuracy		F1		Kappa	
	Sin contexto	Con contexto	Sin contexto	Con contexto	Sin contexto	Con contexto
SVM	76,49%	92,92%	64,51%	91,75%	52,69%	87,21%
MLP	63,09%	90,09%	48,40%	84,44%	29,86%	81,90%

Tabla XX: Resumen de los modelos de clasificación para la tarea de conducción tramo vuelta montaña

En todas las actividades se observa que el desempeño de la clasificación mejora cuando se agregan las variables de contexto. Además, se observa un muy buen desempeño para el clasificador SVM entre 88,71% y 92,92% de accuracy.

D. ANÁLISIS DE LA TAREA

En esta tarea laboral, además de la etiqueta de fatiga por rendimiento, se utilizará un enfoque propuesto por (Vicente et al, 2016) para detectar el estado del conductor. En este enfoque se utiliza la métrica LF como indicador de la actividad del sistema nervioso simpático y HF como el indicador de la actividad del sistema nervioso parasimpático, siendo estos dos los que componen el sistema nervioso autónomo (SNA). Este constructo se basa en que los estados de vigilia se caracterizan por un aumento de la actividad simpática y/o disminución de la actividad parasimpática, mientras que los estados de relajación y somnolencia se caracteriza por un aumento de la actividad parasimpática y/o una disminución de la actividad simpática (Baharav, 1995). De esta manera es posible detectar desde el corazón y de forma no invasiva los estados de Awake (despierto), Stressed (estresado), Fatigued (fatigado) y Drowsy (Somnoliento) (Figura 23), en que se encuentra el conductor mientras realiza su actividad y desde aquí en adelante esta etiqueta será llamada Fatiga SNA.

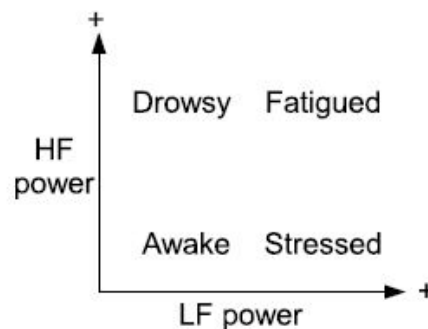


Figura 23: Distribución de estados hipotéticos del sujeto en función de HF y LF de la señal de variabilidad del ritmo cardíaco (Vicente et al, 2016).

9.- Efecto de la tarea por individuo

El análisis comienza utilizando los datos de características de las señales psicofisiológicas obtenidas en el paso anterior para calcular las componentes principales para cada subactividad obteniendo como resultado que el 100% de la variabilidad de las subcaracterísticas normalizadas es explicada por 22 componentes principales para el viaje de ida y 19 para el viaje de vuelta (Anexo II).

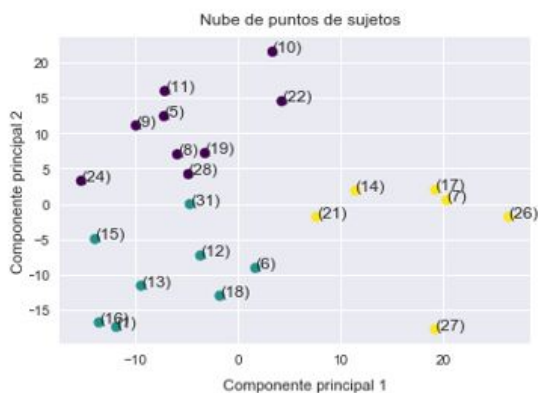
A continuación, se prueban múltiples kernels de clusterización con distinto número de clusters para seleccionar la configuración óptima del modelo de agrupación GMM y se computa los valores del BIC para las distintas configuraciones del GMM para cada subactividad (Anexo III). Además se calcula el silhouette value (Anexo IV) para corroborar el cálculo mediante la convergencia de estos dos métodos, obteniendo los resultados de la Tabla XXI.

Viaje	Trecho	Número de grupos	Kernel
Ida	Ciudad	3	Full
Ida	Montaña	3	Full
Vuelta	Montaña	4	Full
Vuelta	Cuidad	3	Full

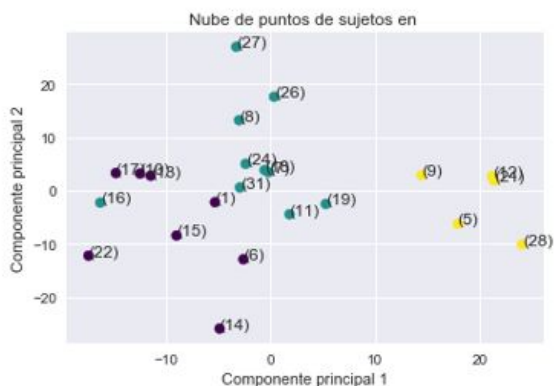
Tabla XXI: Resumen de los modelos de clusterización de individuos.

Con la configuración del kernell full y el número de grupos se procede a definir los clusters propiamente tal. Para esto se ejecutan 10 iteraciones del modelo GMM (Kernel full y 3 clusters) y se descartan las iteraciones donde resultan agrupaciones demasiado desbalanceadas. Luego, se escoge la que posee menor valor del BIC (Anexo V). Utilizando las dos componentes principales de mayor explicación de la varianza, en la Figura 24, se puede visualizar los clusters resultantes para casa subactividad. Finalmente estos grupos se validan realizando una prueba F para determinar la diferencia significativa entre las medias de las señales psicofisiológicas y la prueba de Levene para probar la diferencia significativa de sus desviaciones estándar con respecto a la desviación de todos los sujetos juntos (Anexo VI, VII, VIII, IX, X, XI, XII, XIII).

a) Viaje de ida, trecho ciudad



b) Viaje de ida, trecho montaña



c) Viaje de vuelta, trecho montaña

d) Viaje de vuelta, trecho ciudad



Figura 24: Visualización de los grupos de sujetos usando las dos componentes principales con mayor poder explicativo.

Estos resultados, hacen ver que el kernel full es la mejor opción para agrupar sujetos en función de las componentes principales del comportamiento psicofisiológico durante la ejecución de las tareas. También se observa, en el caso del viaje de regreso es cuando se experimenta una mayor varianza inter sujetos en el comportamiento psicofisiológico determinando que uno de los grupos en el trecho de ciudad y montaña sean un par de sujetos con comportamiento anómalo con respecto al promedio, para la medición realizada. Con estos grupos es posible estudiar la diferencia en su comportamiento de fatiga para poder focalizar las actividades preventivas o bien para testear si es posible realizar clasificaciones y predicciones de la fatiga más precisas dentro de un grupo de similar comportamiento.

10.- Descripción de la tarea para la muestra agregada

En esta etapa se analiza la respuestas comunes de los usuarios a la tarea experimental. Para esto se examinan las hipótesis descritas en la metodología general (punto 1.D). la interpretación de los resultados de los valores p se hará en función de la tabla en el Anexo XIV. A continuación se describe esta exploración.

1. H1. Existe dependencia estadística entre el nivel de fatiga y los grupos de sujetos experimentales.

En esta etapa se analiza la relación existente entre los grupos de sujetos encontrados en la sección 9 con las etiquetas de fatiga SNA y la fatiga de rendimiento. Para esto se genera la tabla de contingencia de cada subactividad donde se especifica el número de etiquetas de fatiga SNA y de rendimiento para grupo de sujetos (Anexo XV y Anexo XVI).

Con los datos de las tablas de contingencia se realiza la prueba de chi-cuadrado para determinar si al comparar entre los grupos de sujetos, existe una diferencia significativa entre las distribuciones de las dos etiquetas de fatiga, esto se repite para cada sub actividad. los valores p resultantes se resumen en la Tabla XXII y el resultado de la visualización de las distribuciones de fatiga en las figuras 25 y 26.

p-valor prueba chi-cuadrada		
Subactividad	Fatiga SNA	Fatiga Rendimiento
IDA CUIDAD	0.000	0.000
IDA MONTAÑA	0.110	0.000
VUELTA MONTAÑA	0.000	0.006
VUELTA CUIDAD	0.000	0.000

Tabla XXII: Resumen de los valores p de las pruebas chi cuadrado a las tablas de contingencia de cada subactividad y etiqueta de fatiga.

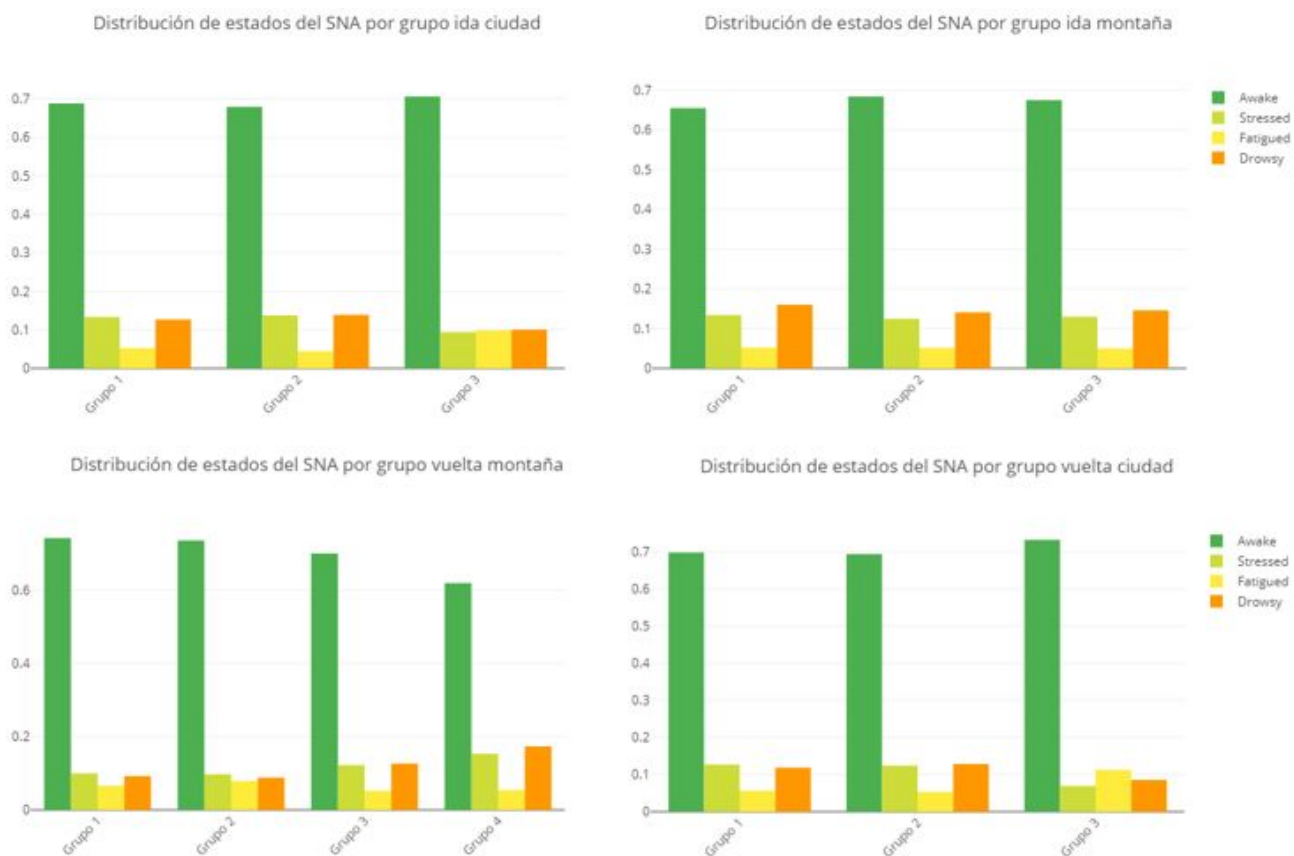


Figura 25: Distribución de niveles de fatiga por grupo considerando la etiqueta de fatiga SNA.

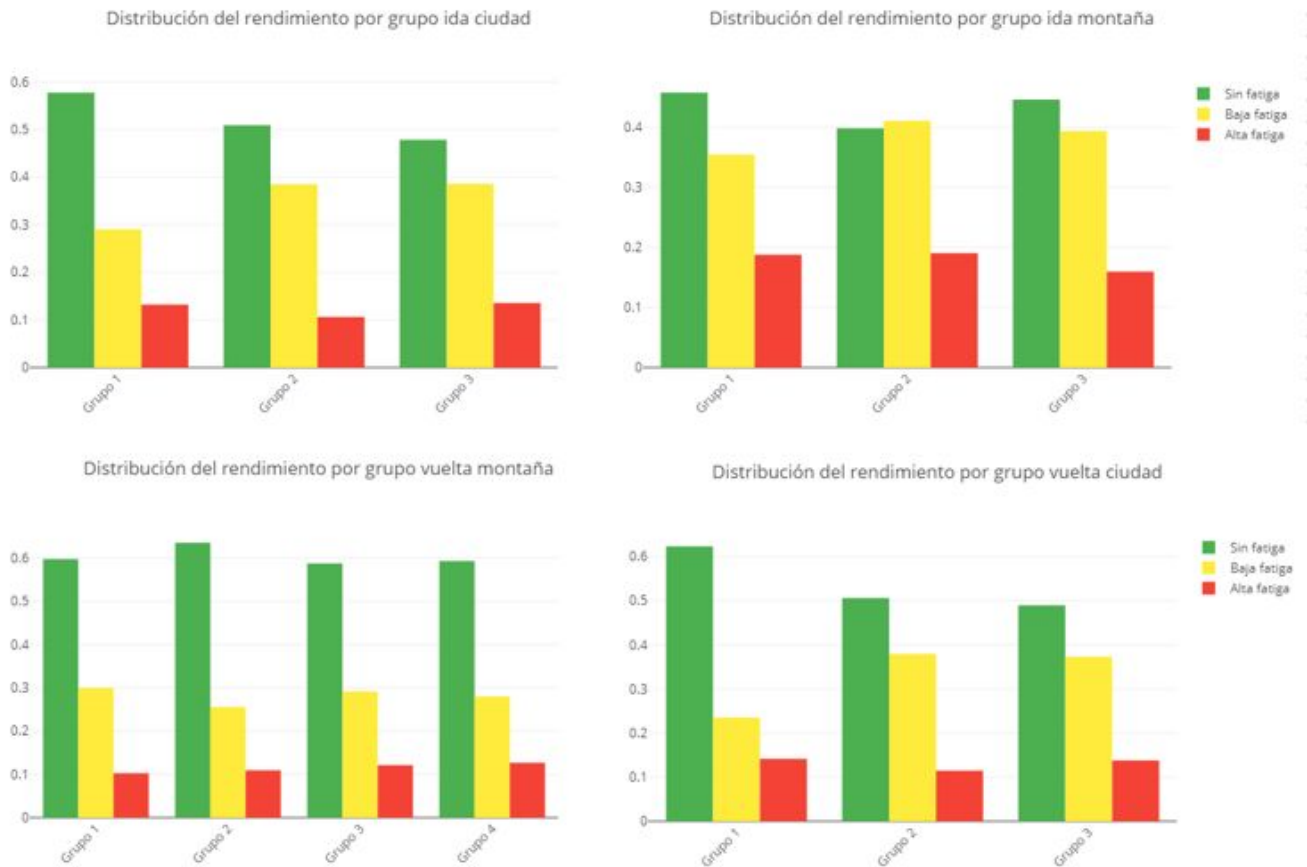


Figura 26: Distribución de niveles de fatiga por grupo considerando la etiqueta de fatiga por rendimiento.

2. H2. Existe una tendencia creciente de la fatiga con el avance temporal de la tarea laboral.

Con esta hipótesis se busca estudiar la evolución de la fatiga en el transcurso de la tarea experimental. Para este análisis se divide el largo temporal de la subactividad en 10 tramos y para cada uno se calcula la fatiga promedio experimentada por los sujetos, en el caso de la fatiga SNA se le asignan los siguientes valores para cada estado, awake: 0, stressed: 1, fatigued: 2, drowsy: 3, que representan un indicador de la pérdida de la vigilia del conductor, mientras que en fatiga de rendimiento, sin fatiga: 0, baja fatiga: 1, alta fatiga: 2, que indica el nivel de calidad de conducción.

Para este caso se observa que la monotonía de las curvas no es constante por lo tanto se analiza la tendencia por tramos (Figura 27) y además se aprovecha de calcular el coeficiente de correlación entre ambas curvas para estudiar la relación entre los estados del SNA y el rendimiento

en la conducción que experimentaron los choferes, cuyos resultados se muestran en la Tabla XXIII. Adicionalmente, se realiza una visualización de los niveles de fatiga georeferenciados en el trecho de montaña (Figura 28 y 29) para estudiar cómo las características de los tramos de la ruta pueden afectar al desarrollo de fatiga y por lo tanto a la seguridad de la operación.

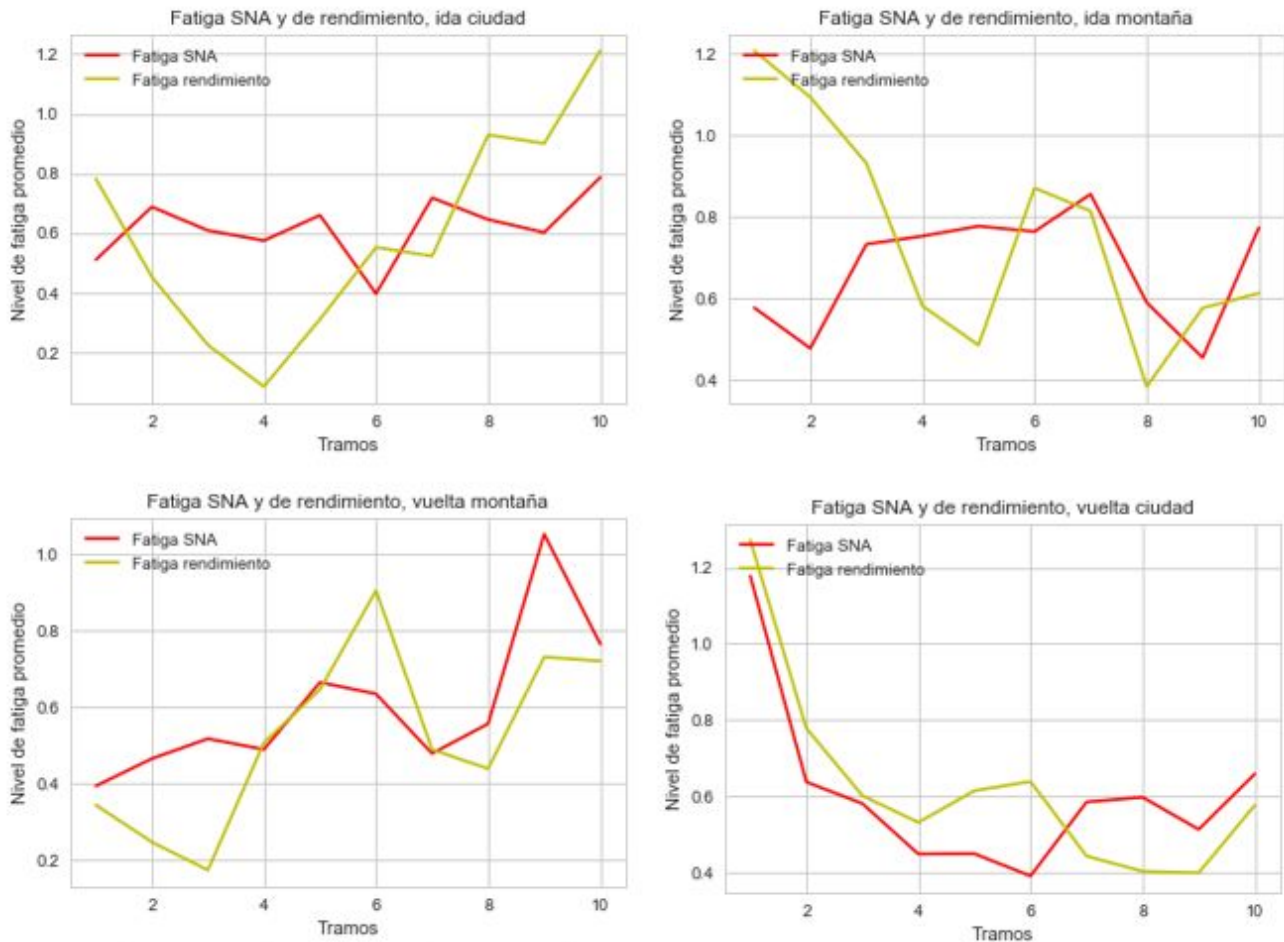


Figura 27: Fatiga promedio por tramos.

Subactividad		Correlación	p-valor
IDA	CIUDAD	0.17	0.62
IDA	MONTAÑA	-0.19	0.57
VUELTA	MONTAÑA	0.73	0.01
VUELTA	CIUDAD	0.21	0.55

Tabla XXIII: Resultados del coeficiente de correlación de Spearman.



Figura 28: Etiqueta de fatiga por rendimiento promedio para montaña ida.



Figura 29: Etiqueta de fatiga por rendimiento promedio para montaña vuelta.

3. H3. Existe diferencia significativa en las medianas y varianzas de las características psicofisiológicas entre los niveles de fatiga.

En esta etapa se busca estudiar los valores promedio y desv. estándar de las características psicofisiológicas normalizadas cuando los sujetos se encuentran en los diferentes niveles de fatiga SNA (despierto, estresado, fatigado, y somnoliento) y de rendimiento (sin fatiga, baja fatiga y alta fatiga). Esto se testeó utilizando la prueba de Kruskal Wallis (Tablas XXIV y XXV) y la prueba de Levene (Anexo XVII y XVIII) y se visualizan las tendencias de los valores promedio y desv. estándar entre los estados de las etiquetas de fatiga ordenados de menor a mayor nivel de vigilia asociado a este estado (Ejemplo Anexo XIX).

Características	Kruskal Wallis Fatiga SNA							
	IDA CIUDAD		IDA MONTAÑA		VUELTA CIUDAD		VUELTA MONTAÑA	
	Tendencia	p-valor	Tendencia	p-valor	Tendencia	p-valor	Tendencia	p-valor
meanHR		0.051		0.003		0.991		0.217
stdHR		0.000		0.000		0.000		0.004
rangeHR		0.000		0.000		0.000		0.030
meanHRV		0.000		0.000		0.000		0.000
stdHRV		0.000		0.000		0.000		0.000
rmssdHRV		0.000		0.000		0.000		0.000
VLF		0.000		0.000		0.000		0.000
TP		0.000		0.000		0.000		0.000
ratioLFHF		0.000		0.000		0.000		0.000
meanEDA		0.000		0.171		0.000		0.000
stdEDA		0.096		0.215		0.000		0.000
meanSCL		0.000		0.001		0.000		0.000
meanSCR		0.000		0.020		0.001		0.153
rangeSCL		0.001		0.000		0.000		0.000
maxSCR		0.000		0.016		0.000		0.022
meanSKT		0.000		0.000		0.005		0.000
stdSKT		0.000		0.566		0.004		0.919
rangeSKT		0.000		0.620		0.002		0.771
Mean_humedad		0.000		0.000		0.000		0.000
std_humedad		0.000		0.000		0.001		0.000
Mean_temp_DFT		0.000		0.065		0.798		0.001
std_temp_DFT		0.000		0.001		0.206		0.065
Mean_temp_CSS		0.000		0.018		0.711		0.027
std_temp_CSS		0.000		0.158		0.000		0.074
Mean_co2		0.000		0.000		0.000		0.000
std_co2		0.000		0.001		0.000		0.001
Mean_sonido		0.521		0.000		0.001		0.000
std_sonido		0.000		0.000		0.000		0.000
mean_acc		0.000		0.004		0.000		0.000
Min_speed		0.008		0.065		0.004		0.000
Max_speed		0.899		0.000		0.068		0.000
Mean_speed		0.395		0.677		0.022		0.000
std_speed		0.000		0.000		0.081		0.000
Fe		0.000		0.001		0.000		0.000
Ne		0.000		0.005		0.000		0.000

Tabla XXIV: Resultados de la aplicación del Test de kruskal Wallis entre los estados de fatiga SNA.

Características	Kruskal Wallis Fatiga rendimiento							
	IDA CIUDAD		IDA MONTAÑA		VUELTA CIUDAD		VUELTA MONTAÑA	
	Tendencia	p-valor	Tendencia	p-valor	Tendencia	p-valor	Tendencia	p-valor
meanHR		0.000		0.000		0.000		0.090
stdHR		0.000		0.000		0.000		0.402
rangeHR		0.000		0.000		0.000		0.115
meanHRV		0.000		0.000		0.000		0.000
stdHRV		0.000		0.365		0.000		0.000
rmsdHRV		0.000		0.000		0.000		0.000
VLF		0.002		0.000		0.000		0.000
TP		0.000		0.000		0.000		0.000
ratioLFHF		0.777		0.447		0.697		0.231
LF		0.000		0.000		0.000		0.000
HF		0.000		0.000		0.000		0.000
meanEDA		0.000		0.000		0.000		0.018
stdEDA		0.000		0.000		0.001		0.013
meanSCL		0.000		0.000		0.000		0.000
meanSCR		0.000		0.000		0.000		0.001
rangeSCL		0.000		0.000		0.005		0.000
maxSCR		0.000		0.000		0.000		0.942
meanSKT		0.000		0.000		0.000		0.000
stdSKT		0.000		0.000		0.000		0.000
rangeSKT		0.000		0.000		0.001		0.000
Mean_humedad		0.000		0.000		0.000		0.000
std_humedad		0.000		0.000		0.000		0.000
Mean_temp_DFT		0.000		0.000		0.640		0.000
std_temp_DFT		0.000		0.000		0.000		0.000
Mean_temp_CSS		0.000		0.000		0.083		0.000
std_temp_CSS		0.000		0.000		0.000		0.000
Mean_co2		0.000		0.000		0.000		0.000
std_co2		0.000		0.000		0.000		0.000
Mean_sonido		0.000		0.000		0.000		0.000
std_sonido		0.000		0.000		0.000		0.000
mean_acc		0.000		0.000		0.000		0.000
Min_speed		0.000		0.000		0.000		0.000
Max_speed		0.000		0.000		0.000		0.000
Mean_speed		0.000		0.000		0.000		0.000
std_speed		0.000		0.000		0.000		0.000

Tabla XXV: Resultados de la aplicación del Test de kruskal Wallis entre los estados de fatiga de rendimiento

11.- Resultados y conclusiones

1. H1. Existe dependencia estadística entre el nivel de fatiga y los grupos de sujetos experimentales.

Los resultados indican que para la mayoría de los casos, la tarea de conducción existe dependencia muy significativa (< 0.01) entre las variables de niveles de fatiga y los grupos de sujetos según su comportamiento psicofisiológico. En los casos de viaje de ida, trecho montaña, no existe diferencia en la fatiga SNA experimentada por los grupos. En las figuras 25 y 26 se aprecian las diferencia entre las distribuciones de niveles de fatiga por grupos de sujetos. Con lo que se puede identificar que hay grupos que experimentan más alta fatiga y por lo tanto podrían ser el foco en actividades preventivas.

2. H2. Existe una tendencia creciente de la fatiga con el avance temporal de la tarea laboral.

Los resultados indican que la tarea de conducción provoca una fatiga con tendencia variable en el tiempo, debido a que al ser una tarea donde los estímulos varían en función del contexto, por ejemplo entre la conducción urbana y una carretera monótona. En el viaje de ida se observa mayor fatiga ligeramente mayor en la segunda mitad del viaje, vale decir en el trecho de montaña, mientras que el rendimiento se mantuvo oscilante alrededor de su valor promedio alcanzando el peor rendimiento (máxima fatiga) en el comienzo de la ruta de montaña.

Por otro lado, en el viaje de vuelta, ambas métricas de fatiga tuvieron un comportamiento similar, obteniendo un 73% de correlación en el trecho de montaña, con significancia estadística y un 21% sin significancia. Cuando comienza la tarea se inicia con una fatiga menor a la experimentada cuando termina el viaje de ida, por lo tanto los sujetos logran recuperarse durante su descanso. En promedio los sujetos viven una fatiga creciente hasta comienzan a conducir en ciudad donde la carga mental de la tarea de conducción es mayor debido a que en ambientes urbanos existen mayor cantidad de demandas y estímulos y provocan una tendencia negativa en la fatiga pero manteniéndose mayor al comienzo debido a la acumulación de carga mental.

3. H3. Existe diferencia significativa en las medianas y varianzas de las características psicofisiológicas entre los niveles de fatiga.

En el caso del análisis con la etiqueta de fatiga SNA, los resultados señalan que cuando los conductores se encuentran en un estado de fatiga y/o somnolencia, hay mayores valores en las características que son un indicador de variabilidad del ritmo cardiaco, tales como stdHR, RangeHR, MeanHRV, stdHRV, rssidHRV por el lado de las características de dominio temporal, en las de dominio de frecuencias, VLF toma mayor valor en estado de fatiga, TP se observan mayores valores

cuando el sujeto se encuentra en los estados de fatiga y estrés, mientras que Ratio LF/HF aumenta en el estado de estrés. Por otro lado, en el caso de las características extraídas a partir de la señal de EDA no se observa un patrón repetitivo entre las subactividades para poder concluir de forma generalizada, tanto para la fatiga SNA y de rendimiento, esto se puede deber a las dificultades que presentó el preprocesamiento de la señal de EDA.

Las características calculadas a partir de la señal de temperatura en la piel indica que en la mayoría de los casos se vive un aumento en MeanSKT cuando los sujetos se encuentran en estado de somnolencia y un aumento en la stdSKT cuando los sujetos se encuentran en un estado de fatiga, lo que señala una activación del sistema termorregulador. Por otro lado, se observa mayores valores en el Mean_acc, frecuencia de errores (Fe) y gravedad de los errores (Ne) cuando los sujetos están en estados de fatiga y somnolencia, lo que confirma la estrecha relación entre los niveles de vigilia dictada por los estados del sistema nervioso autónomo y la calidad de la ejecución de la conducción.

Los resultados de la prueba de kruskal wallis al comparar las distribuciones características entre los estados definidos por el indicador de fatiga de rendimiento muestran que existe un aumento muy significativo de los valores de las características de dominio temporal, indicadoras de la variabilidad del ritmo cardiaco cuando los sujetos se encuentran en los estados de baja fatiga y/o alta fatiga, mientras que los valores más bajos se registraron cuando el sujeto no experimentaba fatiga, lo que confirma la teoría que enlaza el HRV con la fatiga y somnolencia (Malik, et al, 2018). También, las características de dominio de frecuencias del HRV indican un aumento en los estados de fatiga, a excepción del Ratio LF/HF que disminuye junto con el desarrollo de la fatiga, coincidiendo con los resultados del mismo análisis con la etiqueta de fatiga SNA.

Los análisis sobre ambas etiquetas también coinciden con el aumento de la MeanSKT cuando aumenta la fatiga, sin embargo difieren en que la fatiga de rendimiento aumenta cuando disminuyen los indicadores de variabilidad de la temperatura en la piel (stdSKT y rangeSKT). Además otro resultado consistente es que dentro de la cabina del bus existió un aumento en el nivel de ruido promedio y su variabilidad, cuando la conducción se ejecutó con peor rendimiento.

4.- ANÁLISIS COMPARATIVO DE LA APLICACIÓN DE LA METODOLOGÍA

En este capítulo se realiza un análisis comparativo de la aplicación de la metodología expuesta en el capítulo 2 en base a tres actividades donde se ha puesto en práctica, dos en un ambiente controlado de laboratorio: digitación y reposición, cuyo análisis completo va en documentos adjuntos; y una realizada en terreno en un ambiente real: conducción de buses. En particular, se exponen los puntos principales de cada bloque de la metodología en tablas comparativas para las tres actividades y su correspondiente discusión. A continuación, se presenta una breve descripción de cada tarea:

- Digitación:** Esta labor consiste en digitar los campos de una serie de facturas en un formulario a través de una aplicación computacional. Tal como se muestra en la figura 30, esta labor es netamente repetitiva y se compone de sólo una actividad - la digitación - y una decisión.

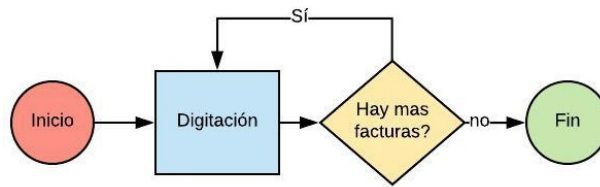


Figura 30: Representación esquemática de la tarea de digitación

- Reposición:** Esta labor consiste en la colocación de un conjunto de cajas en un estante etiquetado de acuerdo a un instructivo. Tal como se muestra en la figura 31, esta labor es repetitiva y se compone de 3 actividades - lectura y organización de las cajas y reporte - y una decisión.

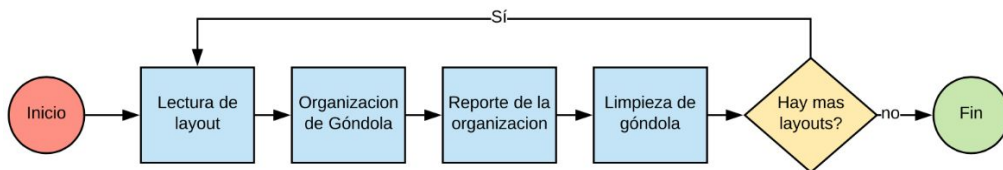


Figura 31: Representación esquemática de la tarea de reposición

- **Conducción:** Labor de conducción profesional de buses de traslado de pasajeros a la mina Los Bronces. Como es propio de una tarea de conducción requiere mantener la atención en un estímulo constante.

A.- DEFINICIÓN DE LA TAREA

1.- Caracterización de la tarea

La tabla XXVI muestra la descripción de las tareas analizadas según las dimensiones de “Monotonía”, “Repetitividad” y “Continuidad” explicadas en la metodología (Sección 2).

Tarea/Hito	Actividades	Monotonía	Repetitividad	Continuidad
Digitación: Digitar una factura en una planilla digital.	<ul style="list-style-type: none"> ● Leer un documento físico ● Escribir en un formulario digital ● Buscar un nuevo documento 	<i>Monotona</i>	<i>Repetitiva</i>	<i>Continua</i>
Reposición: Organización de cajas según layout y reporte del resultados.	<ul style="list-style-type: none"> ● Leer un instructivo en pantalla ● <i>Manipular cajas según instrucciones</i> ● <i>Fotografiar el resultado</i> ● <i>Limpiar las góndolas</i> 	<i>No Monotona</i>	<i>Repetitiva</i>	<i>Continua</i>
Conducción: Conducir bus con pasajeros entre Santiago y Mina Los Bronces.	Todo lo que involucra una conducción profesional en camino de ciudad y montaña	<i>No Monotona</i>	<i>No Repetitiva</i>	<i>Continua</i>

Tabla XXVI: Caracterización de cada tarea laboral.

En cuanto a las métricas de rendimiento, se observa en la tabla XXVII que para las tareas en entorno controlado se definen métricas similares de rendimiento, mientras que para la tarea de conducción se utilizan métricas distintas debido a las diferencias de las condiciones de ejecución.

Actividad	Métrica de efectividad	Métrica de eficiencia
Digitación	Errores por factura = (Número total de campos incorrectos por factura)	Frecuencia de errores = Cantidad de veces que al menos un error en un documento Tiempo medio de ejecución = tiempo que demoró en promedio en completar la digitación de un documento

		Varianza del tiempo de ejecución = Establece cambios al ritmo promedio de ejecución de la tarea
Reposición	Errores por repetición = (Número de cajas en ubicación incorrectos) + (Número de errores en secuencias en plataforma computacional)	Frecuencia de errores = Cantidad de veces que se equivocó al menos una vez Tiempo medio de ejecución = tiempo que demoró en promedio en completar un ciclo de actividades Varianza del tiempo de ejecución = Establece cambios al ritmo promedio de ejecución de la tarea
Conducción	Aceleraciones normalizadas sobre 0,5	Cantidad de veces que la aceleración normalizada obtuvo valor superior a 0,5, durante la ventana de tiempo.

Tabla XXVII: Métricas de rendimiento para cada tarea laboral.

2.- Caracterización del entorno de ejecución

La tabla XXVIII muestra el resumen de la caracterización del entorno directo e indirecto para cada tarea. Es claro que para actividades laborales reales el entorno indirecto no se puede controlar ni obviar.

Entorno	Digitación	Reposición	Conducción
Directo	Sala de experimento, condición de laboratorio	Sala de experimento, condición de laboratorio	Cabina del conductor, durante una ruta real. Condición de terreno.
Indirecto	Edificio institucional, sin injerencia en el ambiente	Edificio institucional, sin injerencia en el ambiente	Condiciones de la ruta

Tabla XXVIII: Descripción del entorno de ejecución de las actividades de toma de datos para cada tarea laboral.

3.- Definición de la muestra

La tabla XXIX muestra el grupo experimental para cada tarea. En el caso de conducción se midieron todos los choferes posibles, siendo todos hombres. Para las tareas de digitación y reposición se solicitaron voluntarios, en cambio, para conducción fueron sólo conductores profesionales con varios años de experiencia.

	Digitación	Reposición	Conducción
Participantes	16	16	30
Edad (años)	31,7 ± 8,9	30,8 ± 10,9	47,5 ± 21,3
Género (M/F)	12 / 4	10 / 6	30 / 0

Tabla XXIX: Descripción de la muestra de sujetos experimentales para cada tarea laboral.

4.- Preparación de artefactos y materiales

Para los tres experimentos realizados, se utilizaron los mismos artefactos propuestos en la metodología para cada medida:

- I. **Medidas fisiológicas:** Biomonitor.
- II. **Medidas Contextuales**
 - A. **Contexto de entorno:** sensor de contexto.
 - B. **Características del sujeto experimental:** Cuestionario de identificación personal y escala de autoreporte.
- III. **Formularios legales:** Consentimiento informado.
- IV. **Otros, sólo para los casos de digitación y reposición:** Se utilizó cámaras de video y grabación de pantalla, y se crearon aplicaciones con Visual Basic para los formularios de digitación e instructivos para la reposición.

Respecto al primer bloque de la metodología, la definición de la tarea, es se observa que es fundamental tener un claro entendimiento previo de la actividad a analizar. Gracias a esto se pueden definir métricas importantes como la de rendimiento y ajustar el protocolo experimental a los requerimientos y requisitos propios de la tarea, a fin de evitar problemas durante la ejecución del experimento en sí. Además, se observa la relevancia del entorno indirecto en tareas laborales en terreno que es el enfoque principal donde se desea aplicar esta metodología.

B.- DEFINICIÓN DEL EXPERIMENTO

5.- Definición del protocolo experimental

Los 3 experimentos realizados siguieron el diseño de dividir en 3 etapas principales la definición del protocolo: antes, durante y después. Los pasos definidos en particular para cada uno, se ajustaron a la tarea propuesta.

6.- Ejecución del protocolo experimental

La captura de datos se realizó según lo establecido en cada protocolo, con una duración completa por sesión experimental de entre 3 y 3,5 horas, para cada tarea. No se reportan dificultades en la toma de datos que signifiquen pérdidas de información ni variaciones a los protocolos definidos en las tareas de laboratorio. Sin embargo, para la tarea en terreno fue necesario recurrir a un computador resistente al ambiente de faenas y con suficiente autonomía para que la batería durara todo el experimento. En las Figuras 32, 33 y 34 se observa a participantes durante la ejecución de las tareas laborales de digitación, reposición y conducción, respectivamente.



Figura 32: Sujeto experimental realizando la tarea de digitación.



Figura 33: Sujeto experimental realizando la tarea de reposición.



Figura 34: Sujeto experimental realizando la tarea de conducción.

Los participantes ejecutaron sin problemas los experimentos en laboratorio siguiendo las instrucciones. Para el caso del experimento de conducción, por tratarse de una experiencia en terreno en un ambiente no controlado y dentro del marco de una jornada de trabajo, presentó las situaciones y complejidades propias del día a día, como por ejemplo, factores emocionales del conductor dado que está obligado a participar, a diferencia de los participantes en el laboratorio que fueron de manera voluntaria; también la cantidad de tareas y rutina propias de la actividad laboral, además de las condiciones de la ruta, como el tráfico, que imponía distintas duraciones a cada sesión. En el punto 4.C se explicará los efectos de estos factores en las mediciones.

C.- PREPARACIÓN DE DATOS

7.- Limpieza y procesamiento de datos

Siguiendo la metodología propuesta, los datos obtenidos de los 3 experimentos fueron sometidos a los mismos pasos de limpieza y procesamiento establecidos. En la tabla XXX un resumen con las principales consideraciones a cada caso. En este punto es importante destacar que en la tarea de conducción, al realizarse mediciones en el mundo real, requiere un esfuerzo mayor en el procesamiento que las otras tareas. El algoritmo aplicado para procesar la señal PPG y obtener HR y HRV, explicado en el punto 2.C.7 muestra un buen desempeño al procesar la señal y eliminar el ruido. En cambio, el algoritmo para procesar la señal EDA no siempre logra procesar correctamente la señal ruidosa, por lo que fue necesario realizar, en algunos casos, parte de la limpieza de forma manual por juicio experto. Esto deja el desafío de mejorar este algoritmo para el procesamiento de señales muy ruidosas de EDA.

Por otro lado, para cada tarea se obtuvo un número similar de muestras (aproximadamente diez mil). En el caso de conducción, como se analizó la tarea en 4 subactividades, también se realizó el análisis con ese orden de muestras.

	Digitación	Reposición	Conducción
Tamaño Ventana (minutos)	5	10	5
Deslice (segundos)	10	10	10
Datos eliminados	10%	10%	25%
Muestras efectivas obtenidas	10.128	9.036	43.017

Tabla XXX: Reporte resumen de la ventana de tiempo utilizada y de la cantidad total de muestras obtenidas.

8.- Etiquetado de fatiga

La metodología indica 3 componentes para el etiquetado de fatiga: autoreportada, observación y rendimiento. Para todos los experimentos la fatiga autoreportada fue la escala KSS, donde se solicita a cada individuo indicar, de 1 a 9, cómo se siente respecto a su nivel de alerta. A su vez, para la fatiga observada, se asignaron puntajes a los gestos de los participantes. Esta escala no pudo ser aplicada en conducción debido a la posición del experimentador y falta de cámara de video. Detalles del etiquetado se presentan en la tabla XXXI.

Para validar las respuestas del cuestionario KSS se calcula el estadístico alfa de Cronbach cuyo valor sobre 0,8 indica alta consistencia en las respuestas. Para el caso de los experimentos en laboratorio se encuentra una alta consistencia en las respuestas. En contraste, para la actividad en terreno el valor es de 0,61 lo que no es suficiente para validar las respuestas del cuestionario por lo que se descarta esta etiqueta. La existencia de un sesgo de respuesta de los conductores pudiera explicarse por el hecho de encontrarse en su trabajo temieran represalias por exponer respuestas que denotaran que se sentían fatigados.

	Digitación	Reposición	Conducción
Autoreportada / Consistencia (alfa de Cronbach)	KSS / 0,97 Aplicada	KSS / 0,92 Aplicada	KSS / 0,61 NO Aplicada
Observada	Aplicada	Aplicada	NO Aplicada
Rendimiento	-Indicador de error. -Tiempo medio de ejecución. -Varianza del tiempo de ejecución	-Indicador de error. -Tiempo medio de ejecución. -Varianza del tiempo de ejecución	-Indicador de error en base al análisis de acelerómetro del bus.
Cantidad de clusters de fatiga	3	3	Ida-ciudad: 3 Ida-montaña: 3 Vuelta-montaña: 3 Vuelta-ciudad: 3
Mejor modelo / Accuracy / F1	SVM/ 83,11% / 70,40%	SVM / 87,49% / 85,57%	Ida-ciudad: SVM/89,93%/87,26% Ida-montaña: SVM/88,71%/88,51% Vuelta-montaña: SVM/92,92%/91,75%

Vuelta-ciudad:
SVM/89,77%/87,62%

Tabla XXXI: Resumen de los resultados del etiquetado de la fatiga para cada tarea laboral.

En todos los casos se encontró y validó que la cantidad óptima de clusters para las etiquetas es de tres. Además, de los modelos de clasificación probados el que mejor rendimiento muestra es SVM que da un alto porcentaje de acierto en la clasificación de las tareas de laboratorio considerando las características psicofisiológicas. Por otro lado, para conducción SVM da una clasificación del entre 70% y 75% considerando sólo las características psicofisiológicas. En cambio, al agregar las variables de contexto, el nivel de acierto sube a valores entre 88% y 92%. Se concluye entonces que medir el contexto ayuda a estimar de mejor manera la fatiga de rendimiento en conductores.

D. ANÁLISIS DE LA TAREA

9.- Efecto de la tarea por individuo

Para cada experimento se calculó y validó la cantidad de clusters, en las cuales se agruparon los individuos según su respuesta psicofisiológica para la tarea laboral, dando como resultado la tabla XXXII. Se observa que para la mayoría de los casos se encuentran tres grupos distintos, grupos son validados estadísticamente. Este resultado es interesante porque al existir distintos grupos se puede pensar en que un grupo que percibe más fatiga necesitaría mayor intervención que los otros.

	Digitación	Reposición	Conducción
Cantidad de grupos	3	3	Ida-ciudad: 3 Ida-montaña: 3 Vuelta-montaña: 4 Vuelta-ciudad: 3

Tabla XXXII: Grupos de sujetos experimentales en función de su comportamiento psicofisiológico durante la ejecución de la tarea laboral.

10.- Descripción de la tarea para la muestra agregada

Al analizar la tarea con la muestra agregada, es decir, considerando todos los participantes juntos por tarea laboral, se obtuvieron los siguientes marcadores fisiológicos para cada experimento,

que se presentan en la tabla XXXIII. La señales que se repiten en los tres casos es HR y HRV. Con esto se puede concluir el valor de la medición de PPG durante la tarea laboral para caracterizar su fatiga.

	Digitación	Reposición	Conducción
Marcadores fisiológicos	<ul style="list-style-type: none"> -HR -Desviación estándar del EDA -Media del SCR -Rango del SCL -Máximo del SCR -Número de peaks del SCR -Media del HRV, la media del EDA. 	<ul style="list-style-type: none"> -Rango HR -rmssdHRV -HF -Número de peaks del SCR. 	<ul style="list-style-type: none"> - stdHR - RangeHR - MeanHRV - stdHRV - rmssdHRV - ratioLFHF - VLF - TP - MeanSKT - stdSKT - Mean_sonido - std_sonido - Mean_acc - Ne - Fe

Tabla XXXIII: Resumen de los marcadores fisiológicos de fatiga encontrados para cada tarea laboral.

11.- Resultados y Conclusiones

En la tabla XXXIV se presenta el resumen de las hipótesis testeadas para cada una de las tareas laborales. Como se observa no necesariamente se cumplen todas ni de manera completa. Por ejemplo, en la tarea de digitación y reposición se observa la existencia de grupos estadísticamente significativos con comportamiento psicofisiológico distinto. Para conducción también es así excepto para el caso de la observación de fatiga SNA para ida montaña. Por otro lado, para digitación se observa una clara tendencia creciente, en cambio, para reposición el nivel de fatiga promedio oscila constantemente alrededor de su valor promedio, mientras que en conducción depende de la subactividad de conducción definida por su contexto (ciudad/montaña o viaje de ida/regreso).

	Digitación	Reposición	Conducción
1. H1. Existe dependencia estadística entre el nivel de fatiga y los grupos de sujetos experimentales.	SI	SI	SI*
2. H2. Existe una tendencia creciente de la fatiga con el	SI	NO	SI. Parcialmente

avance temporal de la tarea laboral.			
3. H3. Existe diferencia significativa en las medianas y varianzas de las características psicofisiológicas entre los niveles de fatiga.	SI	SI	SI

Tabla XXXIV: Resumen del análisis descriptivo de las tareas laborales.

Independiente si se validan o no las hipótesis, el testeo de éstas permite cumplir el objetivo de la metodología que es caracterizar la tarea laboral, entendiendo que existen matices, grados de impacto y explicaciones particulares de cada tarea. Además, podemos concluir que estamos validando una metodología replicable a otras tareas, donde es fundamental la identificación de las métricas para etiquetado, las cuales deben ser capaces de identificar e incluir la esencia de la tarea descrita.

5.- ANÁLISIS EXPLORATORIO DE LA PREDICCIÓN DE FATIGA USANDO MÉTODOS DE DEEP LEARNING

Esta sección tiene por objetivo explicar el desarrollo de un prototipo de un servicio de inteligencia artificial que permita predecir en tiempo real el estado del sistema nervioso autónomo en que se puede encontrar un conductor (fatiga SNA), estos son despierto, estresado, fatigado y somnoliento, ordenados de izquierda a derecha en cómo disminuye su capacidad de vigilia. Esta predicción va a permitir anteponerse a que el sujeto se encuentre en un estado cognitivo inseguro y se puede gatillar una acción preventiva en la cabina de buses inteligentes y del conductor o por parte de la la empresa adherente, de IST u otro servicio de prevención correctiva.

El modelo de Deep Learning utilizado es Long short-term memory (LSTM), perteneciente a los modelos de redes neuronales recurrentes, los cuales han demostrado tener grandes capacidades para detectar patrones temporales y resolver problemas de clasificación y regresión con buen performance cuando se trata de una variable dependiente que pertenece a una secuencia o serie de tiempo (Guo, et. al., 2016)(Fu, et. al., 2016).

El problema es abordado como una regresión de las series de tiempo de las características del HRV, LF y HF, los cuales se calculan cada 10 segundos y son las métricas que funcionan como indicador de la activación del sistema nervioso simpático y parasimpático respectivamente, esto se

explica con más detalle en el punto d. análisis de la tarea. Para detectar el estado, las variables de LF y HF fueron normalizadas entre el rango 0 y 1 utilizando MinMax y se fijó el valor 0,5 en ambas variables para dividir los cuadrantes que determinan el estado del sistema nervioso autónomo (Figura 23).

Para entrenar la red LSTM se utiliza la metodología de LOOVC (Leave-one-out cross-validation) con 16 sujetos del viaje de ida con ambos trechos, ciudad y montaña, donde se realiza el entrenamiento de una red de 4 capas LSTM, dejando 1 sujeto como conjunto de prueba y los 15 restantes como conjunto de entrenamiento, realizando 16 iteraciones cambiando el sujeto de prueba sin repetirlo. En cada iteración se le entregan 1 batch de entrenamiento (15 épocas en total) que contiene toda la serie de tiempo LF o HF de un sujeto y se ordena aleatoriamente el orden en que ingresan los sujetos de entrenamiento. El algoritmo de optimización utilizado es Adams y RMSE como función de costo y métrica de evaluación de la regresión.

Luego predicciones de los valores de LF y HF son usados para detectar el estado del SNA y como métrica de evaluación se utiliza el recall del estado de somnolencia debido a que es el estado donde el error tipo I es más costoso (predecir que el sujeto estará en no somnolencia cuando realmente sí lo estará). Con esta configuración de entrenamiento, el mejor recall se obtuvo usando como input solo los 3 valores consecutivos previos del LF y HF, vale decir realizando una autorregresión de estas series de tiempo. Los resultados de los valores de la función de costos vs el número de épocas de entrenamiento de una ejecución del LOOCV se ilustran en la Figura 35, donde se observa un comportamiento esperable de disminución con respecto al avance de épocas, donde el test se mantiene con peor performance que el train excepto cuando en una época ingresa los datos de un sujeto con una señal con comportamiento diferente al promedio o bien con valores atípicos.



Figura 35: Ejemplo de función de costos versus las épocas de entrenamiento, donde cada época corresponde al entrenamiento con todos los datos de un sujeto

A manera de ejemplo se ilustra la capacidad predictiva del modelo al testearlo con las señales LF y HF de sujetos aleatorios (Figura 36).

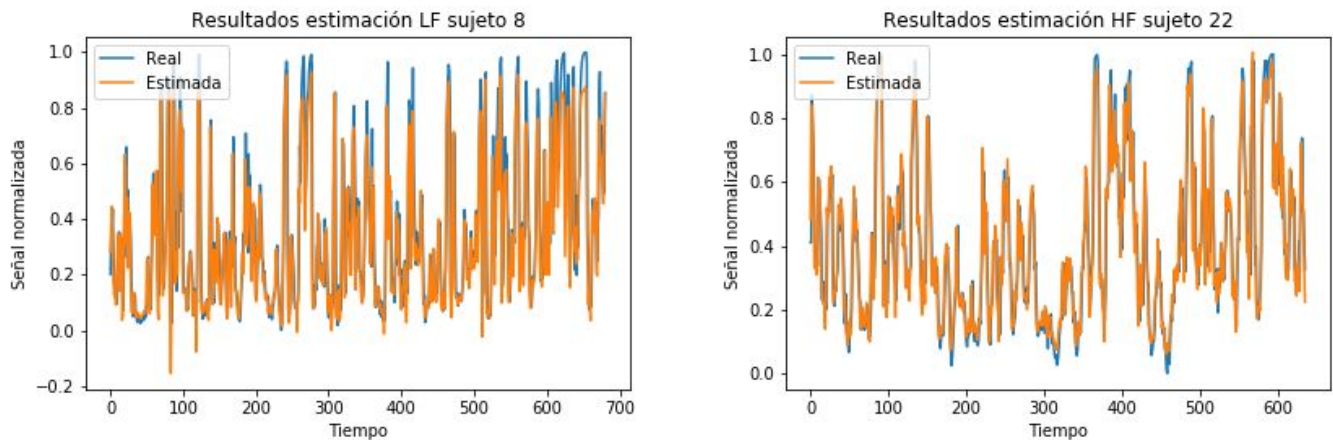


Figura 36: Ejemplos de la visualización del ajuste entre la señal real (azul) y la estimada por el modelo LSTM (naranja)

Con esto se calcula la matriz de confusión con los valores del recall de predicción con 10 segundos de anticipación de las etiquetas de fatiga SNA (Figura 37) donde el valor del recall del estado de somnolencia es del 72%. Este resultado parcial tiene espacios de mejora debido a que el 28% de los casos en que un sujeto promedio este en el estado de somnolencia el sistema indicará que se encuentra en buen estado. Sin embargo, hasta el minuto es un resultado alentador debido a

que se estaría evitando el 72% de los casos y que el 9% de las veces se estaría clasificando como fatigado, el cual es un estado cognitivo cercano al de somnolencia y que se puede notificar al conductor.

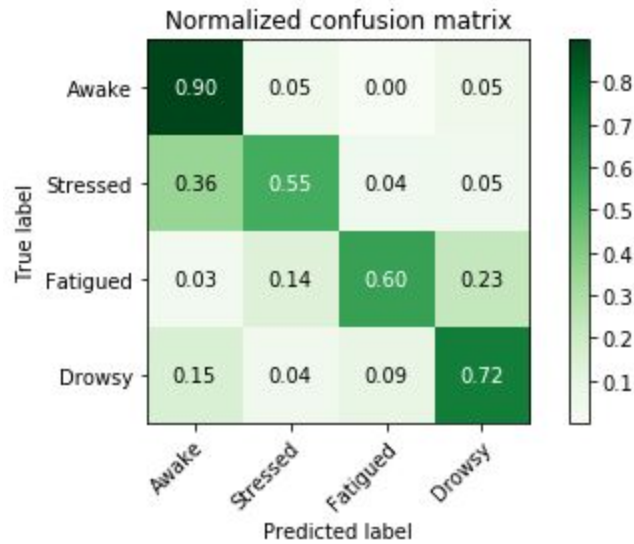


Figura 37: Calculo del recall en la detección de los estados del SNA para cada etiqueta

Además, dado que esta etiqueta de fatiga se puede predecir cada 10 segundos, si se predice un número determinado de estados de somnolencia consecutivamente, aumenta la confianza en la detección del estado inseguro y se puede alertar al conductor. Para mejorar el performance de detección existen múltiples alternativas que se explorarán, tales como modificar la configuración de los datos de entrenamiento, el cómo se ingresan a la red, tamaño del batch y cantidad de épocas de entrenamiento, algoritmo de optimización, función de costos y cantidad de capas LSTM a utilizar, esto con el fin de aumentar la anticipación de la predicción y su recall del estado de somnolencia para así evitar la mayor cantidad posible de accidentes de tránsito por motivos de fatiga mental del conductor.

CONCLUSIONES Y PRÓXIMOS PASOS

En el presente informe se presenta una descripción detallada de la metodología de detección de fatiga aplicada a la ejecución de la tarea de digitación, reposición y conducción profesional.

De la aplicación de la metodología se logra validar su generalidad en entornos controlados y la factibilidad de extraer marcadores psicofisiológicos de fatiga específicos para cada tarea, incluyendo el experimento en terreno. A continuación se detallan las conclusiones y los próximos pasos.

Conclusiones del informe

Las principales conclusiones se detallan a continuación.

- **Se logra construir una metodología completa y replicable:** Se consigue construir una metodología completa para la caracterización de fatiga en tareas laborales que es posible replicar en distintas tareas y con distintos contextos.
- **Validación en terreno de la metodología:** Se logra validar la metodología, antes probada en dos tareas distintas de laboratorio (digitación y reposición), en una tarea laboral real y en terreno (conducción de buses). Las condiciones en terreno presentan mayores desafíos que un experimento en un ambiente controlado, sin embargo, la estructura de la metodología permite generar a priori una correcta definición de los parámetros a medir y un correcto diseño experimental que permite mitigar los problemas propios de una medición en una tarea laboral real como: sesgo de respuesta, ruido de las señales por movimiento y problemas al realizar mediciones de fatiga observada. Además, se encuentra que en las tres tareas distintas existe consistencia en los resultados de etiquetado, clasificación y análisis estadísticos. También es posible encontrar distintos marcadores psicofisiológicos asociados a la tarea y al contexto en que se desenvuelve.
- **Caracterización de tareas laborales:** Es posible caracterizar tareas laborales en función de la evolución de la fatiga de los trabajadores durante la ejecución de una tarea laboral buscando patrones asociados a sus señales psicofisiológicas encontrándose, incluso, grupos de trabajadores con patrones similares. Además, se observa que se debe considerar el contexto y las diferentes componentes de la actividad (como subtareas, estaciones de trabajo, descansos, etc.) que puedan afectar los estados psicofisiológicos de los trabajadores. También, es posible encontrar marcadores psicofisiológicos propios de la tarea

laboral que presentan la superioridad a los instrumentos tradicionales post-tarea de medir en tiempo real la fatiga haciendo posible generar un seguimiento de estos estados.

Próximos pasos

Como se resume en las conclusiones, el trabajo presenta valiosos y auspiciosos resultados en cuanto a la caracterización de la fatiga generada por distintas actividades laborales en tiempo real. A pesar de esto existe aún espacio para mejorar los resultados presentados y extender el trabajo a otras actividades laborales. A continuación se presentan las líneas de investigación y mejoras para el trabajo realizado tanto en el área académica como Ingeniería y/o ergonómica.

1. **Implementación y extensión a otra tareas laborales:** Si bien las experiencias laborales realizadas muestran la generalidad (validez externa) de la metodología presentada, la aplicación a otras versiones de las tareas de conducción, digitación y reposición agregarían información para una mejor definición de su descripción, definición de entornos de ejecución y definición de métricas de ejecución. En esta misma línea es necesario aplicar la metodología a otras actividades laborales para robustecer la metodología.
2. **Implementación de una plataforma web para la ejecución de la metodología:** En la línea de aplicabilidad y facilidad de uso se hace necesario el empaquetamiento web de los artefactos para la realización de los distintos pasos de la metodología y la entrega de reportes en línea sobre la evaluación y evolución de fatiga a nivel agregado (de la tarea sobre los trabajadores) e individual (la evolución de la fatiga durante la realización de la tarea). Este reto en Ingeniería facilitará la extensión de la metodología a otras actividades y su adopción y uso práctico para la administración del riesgo por fatiga.
3. **Investigación en la definición de una tipología de tareas:** Desde el punto de vista de la investigación pura es necesaria una tipología de tareas que incorpore de manera más explícita el conocimiento de ergónomos y psicólogos laborales. Esta definición es importante tanto en la definición de las ventanas de tiempo para la medición de la fatiga como en la definición de métricas de rendimiento. Esto último permitiría mejorar aún más los resultados obtenidos actualmente.
4. **Obtención de más marcadores psicofisiológicos:** Hasta ahora se han utilizado características conocidas de las señales psicofisiológicas. Con la captura de más datos y nuevas tareas laborales es posible la investigación de nuevos marcadores psicofisiológicos y posiblemente de patrones de marcadores según tipología de tareas. Lo anterior permitiría robustecer la aplicación de la metodología y generalizar sus resultados.

BIBLIOGRAFÍA

E. Grandjean, Fatigue in industry, *Occup. Environ.Med.* 36 (1979) 175–186, <http://dx.doi.org/10.1136/oem.36.3.175>.

M.A.S. Boksem, M. Tops, Mental fatigue: costs and benefits, *Brain Res. Rev.* 59 (2008) 125–139, <http://dx.doi.org/10.1016/j.brainresrev.2008.07.001>.

Aryal, A., Ghahramani, A., & Becerik-Gerber, B. (2017). Monitoring fatigue in construction workers using physiological measurements. *Automation in Construction*.

Ryu, K., Myung, R., 2005. Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic. *Int. J. Ind. Ergon.* 35, 991e1009. <http://dx.doi.org/10.1016/j.ergon.2005.04.005>.

M. Feuerstein, et al., *Health Psychology: A Psychobiological Perspective*, Springer, 1986.

Van Daalen, G., Willemsen, T.M., Sanders, K., Van Veldhoven, M.J., 2009. Emotional exhaustion and mental health problems among employees doing "people work": the impact of job demands, job resources and family-to-work conflict. *Int. Arch. Occup.*

Cinaz, B., Arnrich, B., Marca, R.L.A., Tröster, G., 2013. Monitoring of mental workload levels during an everyday life office-work scenario. *Pers. Ubiquitous Comput.* 17, 229e239. <http://dx.doi.org/10.1007/s00779-011-0466-1>.

Kim, Y. K., & Cha, N. H. (2015). Correlations among occupational stress, fatigue, and depression in call center employees in Seoul. *Journal of physical therapy science*, 27(10), 3191-3194.

Cacioppo, J.T., Tassinary, L.G., Berntson, G.G.: *Handbook of Psychophysiology*. Cambridge University Press, Cambridge (2000)

Lin, Y. H., Chen, C. Y., & Lu, S. Y. (2009). Physical discomfort and psychosocial job stress among male and female operators at telecommunication call centers in Taiwan. *Applied ergonomics*, 40(4), 561-568.

Lin, Y. H., Chen, C. Y., Hong, W. H., & Lin, Y. C. (2010). Perceived job stress and health complaints at a bank call center: comparison between inbound and outbound services. *Industrial health*, 48(3), 349-356.

Wegge, J., Vogt, J., & Wecking, C. (2007). Customer-induced stress in call centre work: A comparison of audio-and videoconference. *Journal of occupational and organizational psychology*, 80(4), 693-712.

Hart, S.G., Wickens, C.D., 1990. *Workload Assessment and Prediction*. Manprint. Springer.

Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. *Adv. Psychol.* 52, 139e183.

Goldberg, D. P., & Williams, P. (1988). *A Users' Guide to the General Health Questionnaire*. NferNelson, Basingstoke.

Carlson, D. S., Kacmar, K. M., & Williams, L. J. (2000). Construction and initial validation of a multidimensional measure of work–family conflict. *Journal of Vocational behavior*, 56(2), 249-276.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology*, 54(6), 1063.

Hernandez, J., Morris, R. R., & Picard, R. W. (2011, October). Call center stress recognition with person-specific models. In *International Conference on Affective Computing and Intelligent Interaction* (pp. 125-134). Springer, Berlin, Heidelberg.

Allan, J., Farquharson, B., Choudhary, C., Johnston, D. W., Jones, M. C., & Johnston, M. (2009). Stress in telephone helpline nurses: research protocol for a study of theoretical determinants, physiological aspects and behavioural consequences. *Journal of Advanced Nursing*, 65(10), 2208-2215.

Wilson, G. F., & Russell, C. A. (2003). Real-time assessment of mental workload using psychophysiological measures and artificial neural networks. *Human factors*, 45(4), 635-644.

Cavanaugh, M. A., Boswell, W. R., Roehling, M. V., & Boudreau, J. W. (2000). An empirical examination of self-reported work stress among US managers. *Journal of applied psychology*, 85(1), 65.

Braunwald, "Tratado de Cardiología", Barcelona: Elsevier, 2009.

Sharma, N., & Gedeon, T. (2012). Objective measures, sensors and computational techniques for stress recognition and classification: A survey. *Computer methods and programs in biomedicine*, 108(3), 1287-1301.

J. Zhai, A. Barreto, Stress recognition using non-invasive technology, in: *Proceedings of the 19th International Florida Artificial Intelligence Research Society Conference FLAIRS, 2006*, pp. 395–400.

R.B. Devereux, et al., Left ventricular wall stresses and wall stress-mass-heart rate products in hypertensive patients with electrocardiographic left ventricular hypertrophy: the LIFE study, *Journal of Hypertension* 18 (2000) 1129–1138.

P.E. Bonoris, et al., Significance of changes in R wave amplitude during treadmill stress testing: angiographic correlation, *The American Journal of Cardiology* 41 (1978) 846–851.

D. Novák, et al., EEG and VEP signal processing, Technical Report. Czech Technical University in Prague, Department of Cybernetics, 2004.

R. Horlings, et al., Emotion recognition using brain activity, in: *Proceedings of the 9th International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing*, Gabrovo, Bulgaria, 2008, pp.11.1-1.

W. Liao, et al., A real-time human stress monitoring system using dynamic Bayesian network, in: *Computer Vision and Pattern Recognition—Workshops, CVPR Workshops, 2005*.

L.J.M. Rothkrantz, et al., Voice stress analysis, *Lecture Notes in Computer Science* 3206 (2004) 449–456.

S. Scherer, et al., Emotion recognition from speech: stress experiment, in: Proceedings of the 6th International Language Resources and Evaluation (LREC 2008), Marrakech, Morocco, 2008.

Allan, J., Farquharson, B., Choudhary, C., Johnston, D. W., Jones, M. C., & Johnston, M. (2009). Stress in telephone helpline nurses: research protocol for a study of theoretical determinants, physiological aspects and behavioural consequences. *Journal of Advanced Nursing*, 65(10), 2208-2215.

Mashhadi, M. B., Asadi, E., Eskandari, M., Kiani, S., & Marvasti, F. (2016). Heart rate tracking using wrist-type photoplethysmographic (PPG) signals during physical exercise with simultaneous accelerometry. *IEEE Signal Processing Letters*, 23(2), 227-231.

von Rosenberg, W., Chanwimalueang, T., Adjei, T., Jaffer, U., Goverdovsky, V., & Mandic, D. P. (2017). Resolving ambiguities in the LF/HF ratio: LF-HF scatter plots for the categorization of mental and physical stress from HRV. *Frontiers in physiology*, 8, 360.

Patel, M., Lal, S. K., Kavanagh, D., & Rossiter, P. (2011). Applying neural network analysis on heart rate variability data to assess driver fatigue. *Expert systems with Applications*, 38(6), 7235-7242.

Lilliefors, H. W. (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American statistical Association*, 62(318), 399-402.

Lin, C. T., Chuang, C. H., Huang, C. S., Tsai, S. F., Lu, S. W., Chen, Y. H., & Ko, L. W. (2014). Wireless and wearable EEG system for evaluating driver vigilance. *IEEE Transactions on biomedical circuits and systems*, 8(2), 165-176.

Fu, R., Wang, H., & Zhao, W. (2016). Dynamic driver fatigue detection using hidden Markov model in real driving condition. *Expert Systems with Applications*, 63, 397-411.

Li, G., Lee, B. L., & Chung, W. Y. (2015). Smartwatch-based wearable EEG system for driver drowsiness detection. *IEEE Sensors Journal*, 15(12), 7169-7180.

Lee, Y. C., & Winston, F. K. (2016). Stress induction techniques in a driving simulator and reactions from newly licensed drivers. *Transportation research part F: traffic psychology and behaviour*, 42, 44-55.

Shiwu, L., Linhong, W., Zhifa, Y., Bingkui, J., Feiyan, Q., & Zhongkai, Y. (2011, August). An active driver fatigue identification technique using multiple physiological features. In *Mechatronic Science, Electric Engineering and Computer (MEC), 2011 International Conference on* (pp. 733-737). IEEE.

Morales, J. M., Diaz-Piedra, C., Rieiro, H., Roca-González, J., Romero, S., Catena, A., ... & Di Stasi, L. L. (2017). Monitoring driver fatigue using a single-channel electroencephalographic device: A validation study by gaze-based, driving performance, and subjective data. *Accident Analysis & Prevention*, 109, 62-69.

de Naurois, C. J., Bourdin, C., Stratulat, A., Diaz, E., & Vercher, J. L. (2017). Detection and prediction of driver drowsiness using artificial neural network models. *Accident Analysis & Prevention*.

Lin, C. T., Chuang, C. H., Huang, C. S., Tsai, S. F., Lu, S. W., Chen, Y. H., & Ko, L. W. (2014). Wireless and wearable EEG system for evaluating driver vigilance. *IEEE Transactions on biomedical circuits and systems*, 8(2), 165-176.

Zhang, H., Yan, X., Wu, C., & Qiu, T. (2014). Effect of circadian rhythms and driving duration on fatigue level and driving performance of professional drivers. *Transportation Research Record: Journal of the Transportation Research Board*, (2402), 19-27.

Barwick, F., Arnett, P., & Slobounov, S. (2012). EEG correlates of fatigue during administration of a neuropsychological test battery. *Clinical Neurophysiology*, 123(2), 278-284.

Awais, M., Badruddin, N., & Drieberg, M. (2017). A hybrid approach to detect driver drowsiness utilizing physiological signals to improve system performance and wearability. *Sensors*, 17(9), 1991.

Chuang, C. H., Cao, Z., King, J. T., Wu, B. S., Wang, Y. K., & Lin, C. T. (2018). Brain electrodynamic and hemodynamic signatures against fatigue during driving. *Frontiers in neuroscience*, 12, 181.

Ma, J., Gu, J., Jia, H., Yao, Z., & Chang, R. (2018). The Relationship Between Drivers' Cognitive Fatigue and Speed Variability During Monotonous Daytime Driving. *Frontiers in psychology*, 9, 459.

Dimitrakopoulos, G. N., Kakkos, I., Dai, Z., Wang, H., Sgarbas, K., Thakor, N., ... & Sun, Y. (2018). Functional connectivity analysis of mental fatigue reveals different network topological alterations between driving and vigilance tasks. *IEEE transactions on neural systems and rehabilitation engineering*, 26(4), 740-749.

Sauvet, F., Bougard, C., Coroenne, M., Lely, L., Van Beers, P., Elbaz, M., ... & Chennaoui, M. (2014). In-flight automatic detection of vigilance states using a single EEG channel. *IEEE Trans. Biomed. Engineering*, 61(12), 2840-2847.

Patel, M., Lal, S. K., Kavanagh, D., & Rossiter, P. (2011). Applying neural network analysis on heart rate variability data to assess driver fatigue. *Expert systems with Applications*, 38(6), 7235-7242.

Sahayadhas, A., Sundaraj, K., & Murugappan, M. (2012). Detecting driver drowsiness based on sensors: a review. *Sensors*, 12(12), 16937-16953.

Mohanavelu, K., Lamshe, R., Poonguzhali, S., Adalarasu, K., & Jagannath, M. (2017). Assessment of Human Fatigue during Physical Performance using Physiological Signals: A Review. *Biomedical and Pharmacology Journal*, 10(4), 1887-1896.

Grandjean, E. (1979). Fatigue in industry. *Occupational and Environmental Medicine*, 36(3), 175-186.

May, J. F., & Baldwin, C. L. (2009). Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. *Transportation research part F: traffic psychology and behaviour*, 12(3), 218-224.

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of experimental psychology*, 55(4), 352.

Mehler, B., Reimer, B., & Coughlin, J. F. (2010, September). Physiological reactivity to graded levels of cognitive workload across three age groups: An on-road evaluation. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 54, No. 24, pp. 2062-2066). Sage CA: Los Angeles, CA: SAGE Publications.

Riccio, C. A., Reynolds, C. R., Lowe, P., & Moore, J. J. (2002). The continuous performance test: a window on the neural substrates for attention?. *Archives of clinical neuropsychology*, 17(3), 235-272.

von Rosenberg, W., Chanwimalueang, T., Adjei, T., Jaffer, U., Goverdovsky, V., & Mandic, D. P. (2017). Resolving ambiguities in the LF/HF ratio: LF-HF scatter plots for the categorization of mental and physical stress from HRV. *Frontiers in physiology*, 8, 360.

Michael, R., & Meuter, R. (2006). Sustained attention and hypovigilance: The effect of environmental monotony on continuous task performance and implications for road safety. In *Proceedings of the Australasian road safety research, policing and education conference* (Vol. 10). Monash University.

He, Y., Yan, X., Wu, C., Chu, D., & Peng, L. (2013). Effects of Driver's Unsafe Acceleration Behaviors on Passengers' Comfort for Coach Buses. In *ICTIS 2013: Improving Multimodal Transportation Systems-Information, Safety, and Integration* (pp. 1649-1655).

Dhahir, B., & Hassan, Y. (2018). Studying driving behavior on horizontal curves using naturalistic driving study data. *Transportation research record*, 0361198118784384.

Vicente, J., Laguna, P., Bartra, A., & Bailón, R. (2016). Drowsiness detection using heart rate variability. *Medical & biological engineering & computing*, 54(6), 927-937.

Baharav A, Kotagal S, Gibbons V, Rubin BK, Pratt G, Karin J, Akselrod S (1995) Fluctuations in autonomic nervous activity during sleep displayed by power spectrum analysis of heart rate variability. *Neurology* 45(6):1183–1187

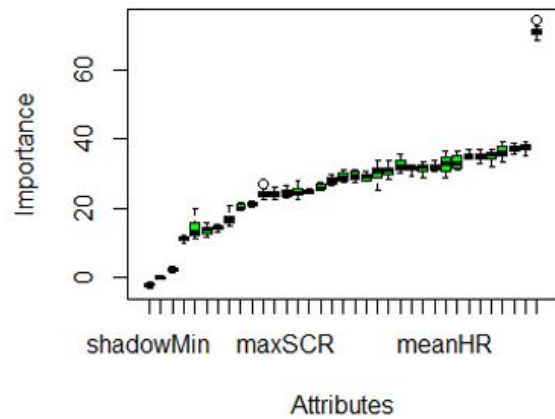
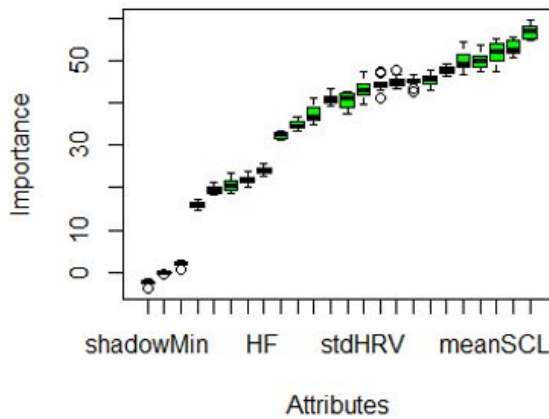
Malik, J., Lo, Y. L., & Wu, H. T. (2018). Sleep-wake classification via quantifying heart rate variability by convolutional neural network. *Physiological measurement*, 39(8), 085004.

Guo, T., Xu, Z., Yao, X., Chen, H., Aberer, K., & Funaya, K. (2016, October). Robust online time series prediction with recurrent neural networks. In *2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA)* (pp. 816-825). Ieee.

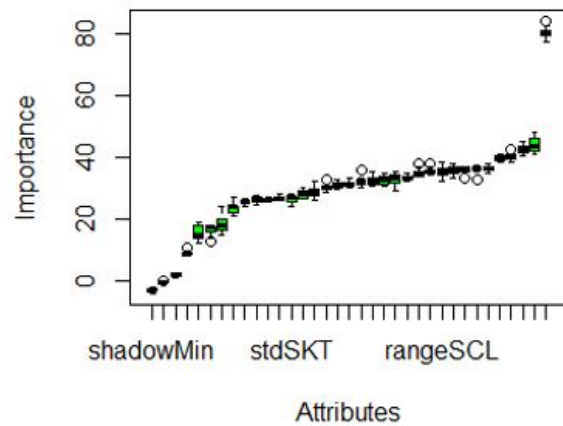
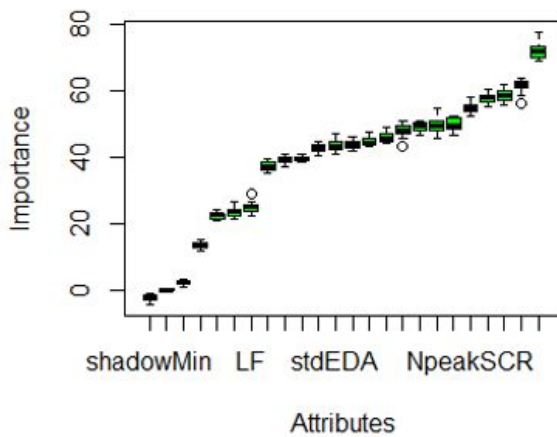
Fu, R., Zhang, Z., & Li, L. (2016, November). Using LSTM and GRU neural network methods for traffic flow prediction. In *2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)* (pp. 324-328). IEEE.

ANEXO I: Gráfico de selección de características con algoritmo Boruta para cada subactividad

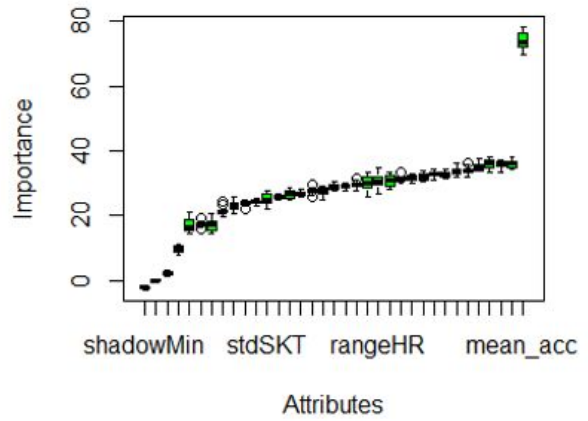
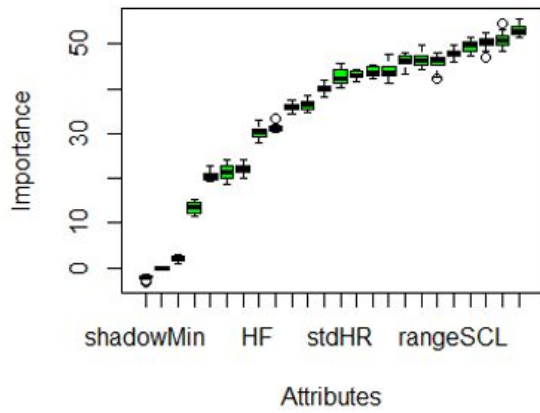
a) Viaje de ida, trecho ciudad. Considerando las variables de contexto y no, respectivamente.



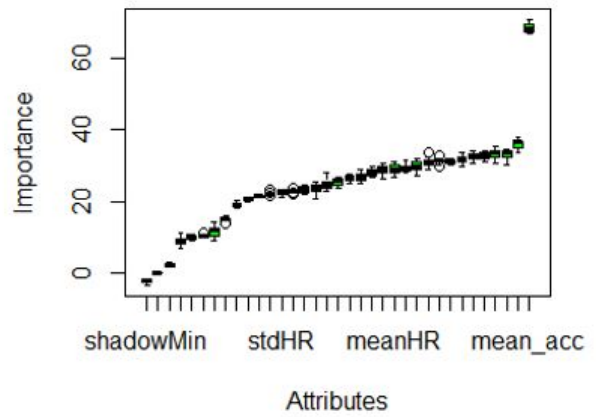
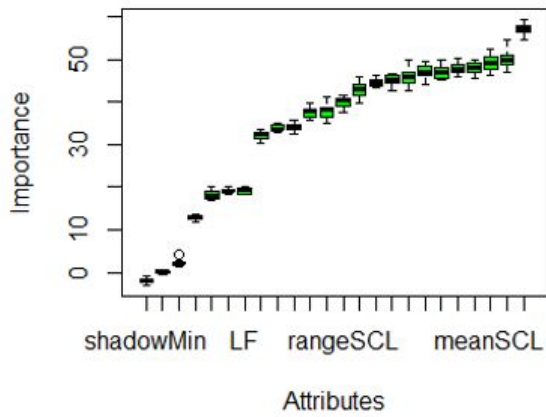
b) Viaje de ida, trecho montaña. Considerando las variables de contexto y no, respectivamente.



c) Viaje de vuelta, trecho ciudad. Considerando las variables de contexto y no, respectivamente.



d) Viaje de vuelta, trecho montaña. Considerando las variables de contexto y no, respectivamente.

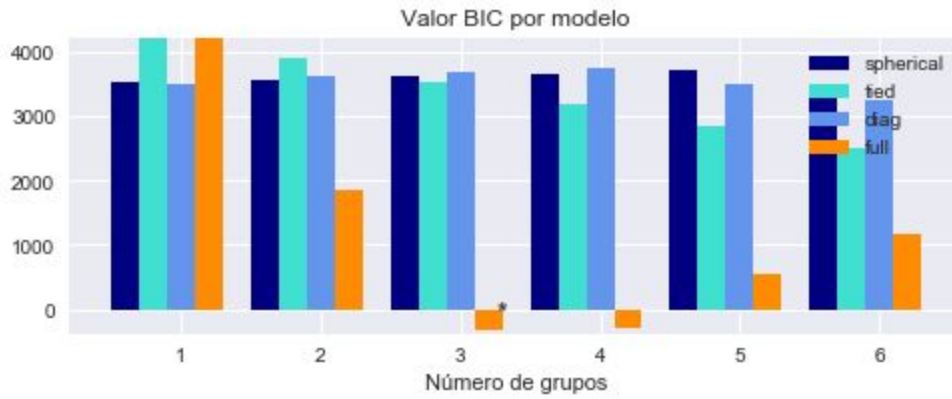


ANEXO II: Tabla de la varianza explicada por las componentes principales para cada subactividad

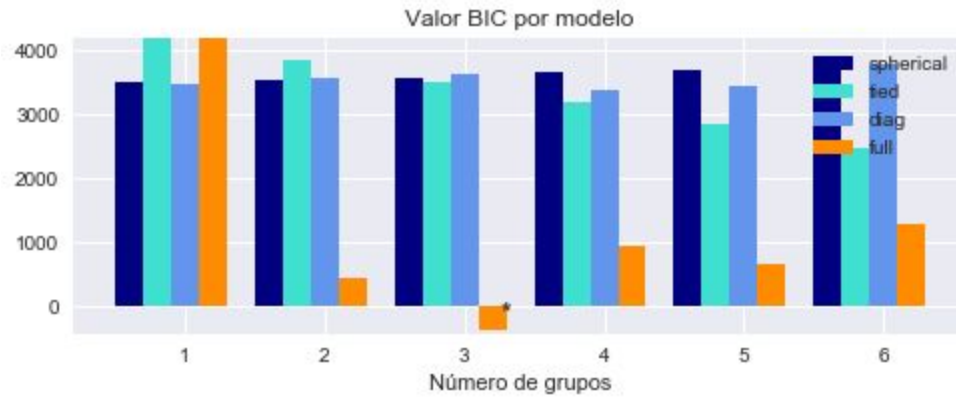
Varianza explicada por las componentes principales (%)										
IDA CIUDAD										
C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
12%	10%	8%	7%	7%	6%	6%	5%	5%	4%	4%
C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22
4%	3%	3%	3%	3%	2%	2%	2%	2%	2%	2%
IDA MONTAÑA										
C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
13%	10%	8%	7%	6%	6%	5%	5%	4%	4%	4%
C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22
4%	3%	3%	3%	3%	3%	2%	2%	2%	2%	2%
VUELTA MONTAÑA										
C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
13%	12%	9%	8%	6%	6%	6%	5%	5%	4%	4%
C12	C13	C14	C15	C16	C17	C18	C19			
4%	3%	3%	3%	3%	2%	2%	2%			
VUELTA CIUDAD										
C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
16%	9%	8%	7%	7%	6%	5%	5%	5%	5%	4%
C12	C13	C14	C15	C16	C17	C18	C19			
4%	4%	3%	3%	3%	3%	2%	2%			

ANEXO III: Valores BIC para decidir kernel y numero de grupos de sujetos, para cada subactividad

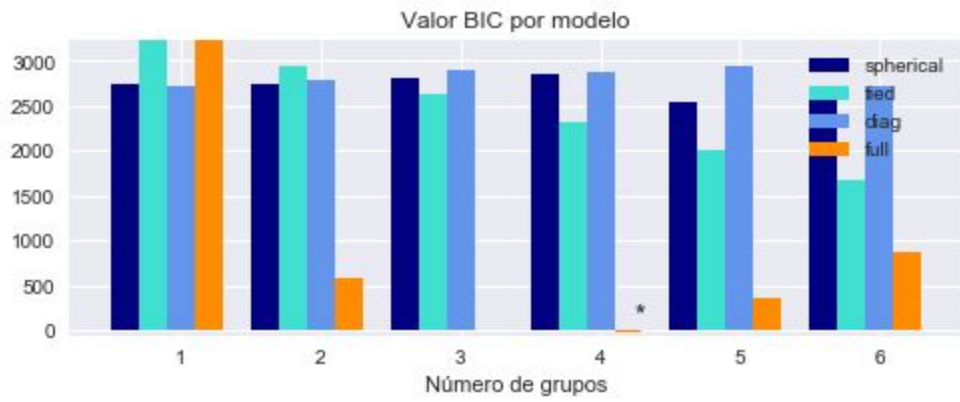
e) Viaje de ida, trecho ciudad



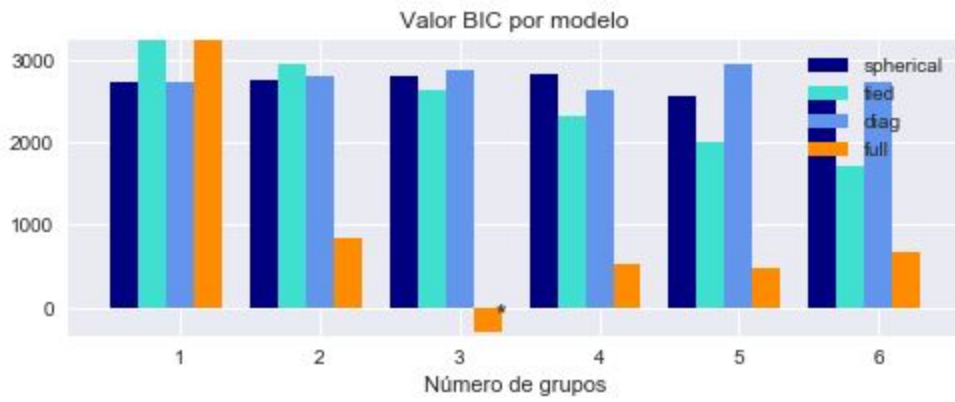
f) Viaje de ida, trecho montaña



g) Viaje de vuelta, trecho montaña

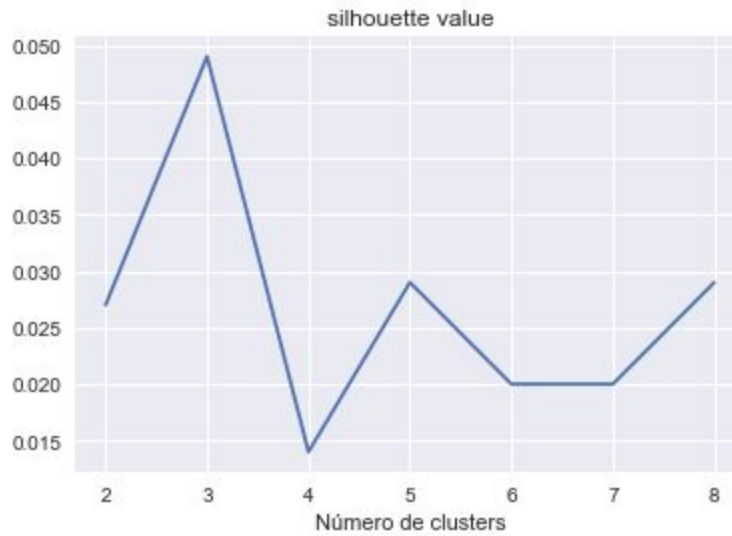


h) Viaje de vuelta, techo ciudad

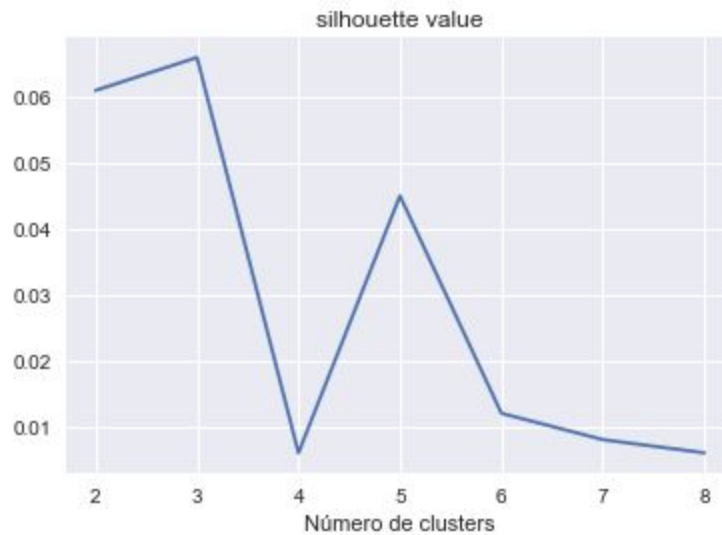


ANEXO IV: Valores del puntaje de Silhouette para determinar la cantidad de grupos de sujetos, para cada subactividad

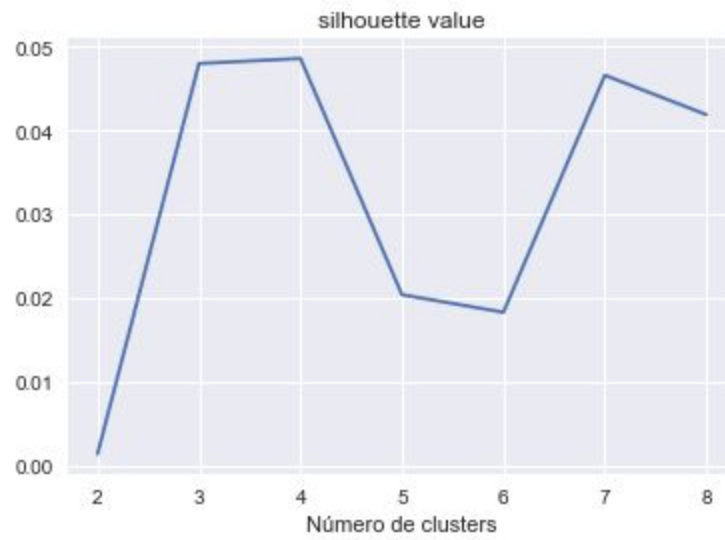
a) Viaje de ida, trecho ciudad



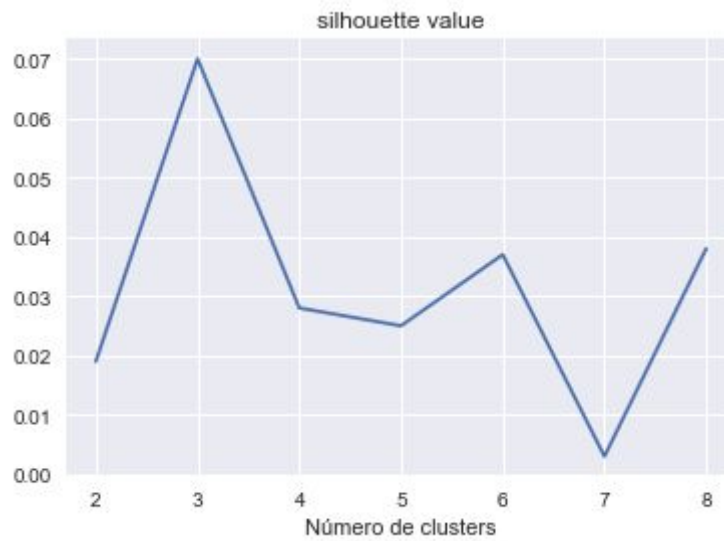
b) Viaje de ida, trecho montaña



c) Viaje de vuelta, trecho montaña



d) Viaje de vuelta, trecho ciudad



ANEXO V: Valores BIC de las iteraciones del modelo de agrupación GMM para cada subactividad

Valores BIC									
1	2	3	4	5	6	7	8	9	10
IDA CIUDAD									
131.0	1497.2	-89.0	-227.2	788.9	506.7	729.9	-306.2	-43.4	-80.7
IDA MONTAÑA									
1051.5	915.6	-309.3	765.7	240.5	1095.0	-346.3	769.6	772.7	344.7
VUELTA MONTAÑA									
-126.9	240.7	25.3	820.3	384.9	2016.8	1147.8	384.4	818.6	
VUELTA CIUDAD									
237.6	389.8	90.3	744.1	821.8	385.2	89.7	565.2	164.1	-198.6

ANEXO VI: Prueba de Levene para la validación de agrupación de sujetos en conducción, ida ciudad

Test de Levene		Tramo 1	Tramo 2	Tramo 3	Tramo 4	Tramo 5	Tramo 6	Tramo 7	Tramo 8	Tramo 9	Tramo 10
HR	std cluster 1	0.904	1.178	1.198	0.880	1.009	0.807	0.830	0.849	0.757	0.915
	p-value cluster 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 2	0.958	0.935	0.998	0.913	0.971	0.975	0.885	0.853	0.912	0.907
	p-value cluster 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 3	1.249	1.072	0.772	0.820	0.712	0.810	0.831	0.889	1.222	1.061
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
std todos los sujetos		1.045	1.125	1.038	0.920	0.935	0.870	0.885	0.877	0.970	1.051
HRV	std cluster 1	0.905	0.922	0.902	0.960	0.930	0.946	0.985	1.024	1.050	1.022
	p-value cluster 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 2	0.922	0.875	0.884	0.951	0.973	1.008	0.978	0.987	0.902	0.946
	p-value cluster 2	0.000	0.000	0.000	0.000	0.000	0.000	0.017	0.000	0.000	0.000
	std cluster 3	1.076	1.235	1.215	1.030	1.171	1.073	0.958	0.998	1.061	0.917
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.047	0.000	0.000
std todos los sujetos		0.963	1.025	1.023	0.979	1.033	1.005	0.976	1.006	1.011	0.976
EDA	std cluster 1	1.193	1.012	1.071	1.177	1.045	0.751	0.472	0.542	0.615	0.596
	p-value cluster 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 2	0.882	0.885	0.890	0.854	1.063	0.876	0.725	0.846	0.695	0.867
	p-value cluster 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 3	1.013	0.810	0.782	0.655	0.782	0.774	0.590	0.578	0.513	0.488
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
std todos los sujetos		1.178	1.040	1.039	1.067	0.992	0.833	0.650	0.721	0.693	0.689
SCR	std cluster 1	2.074	1.293	1.270	1.357	0.647	0.294	0.349	0.330	0.418	0.426
	p-value cluster 1	0.000	0.000	0.169	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 2	0.635	0.742	0.712	0.641	0.908	1.788	0.931	1.256	0.791	0.625
	p-value cluster 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 3	2.139	0.954	0.902	0.713	0.959	0.800	0.605	0.720	1.173	1.027
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
std todos los sujetos		1.786	1.052	1.011	1.019	0.843	1.186	0.680	0.875	0.852	0.767
SCL	std cluster 1	1.146	1.033	0.954	1.192	0.994	0.758	0.445	0.584	0.591	0.570
	p-value cluster 1	0.000	0.000	0.000	0.000	8e-323	0.000	0.000	0.000	0.000	0.000
	std cluster 2	0.639	0.681	0.751	0.726	0.717	0.766	0.723	0.772	0.774	0.814
	p-value cluster 2	0.000	0.000	0.000	0.000	0.000	0.930	0.000	0.000	0.000	0.000
	std cluster 3	1.125	1.027	0.942	0.901	0.908	0.693	0.528	0.534	0.577	0.432
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
std todos los sujetos		1.118	1.042	0.994	1.089	0.889	0.772	0.616	0.689	0.704	0.664
SKT	std cluster 1	1.131	1.062	0.909	0.755	0.815	0.443	0.439	0.444	0.507	0.581
	p-value cluster 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 2	1.008	0.797	0.976	0.890	0.531	0.418	0.460	0.367	0.463	0.564
	p-value cluster 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 3	0.871	0.953	0.854	0.626	0.697	0.981	0.686	0.750	0.647	0.700
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
std todos los sujetos		1.137	0.980	0.984	0.877	0.714	0.635	0.562	0.532	0.570	0.640

ANEXO VII: Prueba de Levene para la validación de agrupación de sujetos en conducción, ida montaña

Test de Levene		Tramo 1	Tramo 2	Tramo 3	Tramo 4	Tramo 5	Tramo 6	Tramo 7	Tramo 8	Tramo 9	Tramo 10
HR	std cluster 1	1.110	0.995	0.750	0.942	0.680	0.738	0.731	0.802	0.924	1.031
	p-value cluster 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 2	0.813	1.138	1.206	0.900	0.847	1.041	0.814	0.850	0.671	0.722
	p-value cluster 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 3	0.966	0.892	1.028	0.969	1.072	1.198	1.189	0.882	0.886	0.922
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
std todos los sujetos		1.096	1.034	1.060	0.941	0.886	0.962	0.923	0.890	0.822	0.898
HRV	std cluster 1	0.912	0.899	0.932	0.969	1.030	1.083	1.053	0.992	1.055	1.116
	p-value cluster 1	0.000	0.000	0.000	0.000	0.006	0.000	0.000	0.000	0.000	0.000
	std cluster 2	0.985	0.965	0.943	0.965	1.012	1.002	0.984	1.034	1.043	1.079
	p-value cluster 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 3	1.057	1.055	1.047	1.041	1.061	1.021	1.020	0.923	0.950	0.992
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
std todos los sujetos		0.993	0.973	0.975	0.996	1.027	1.035	1.011	1.008	1.016	1.076
EDA	std cluster 1	0.604	0.701	0.703	0.866	1.036	0.994	0.709	0.627	0.683	0.522
	p-value cluster 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 2	0.754	0.679	0.650	0.453	0.603	0.546	0.878	0.427	0.871	1.155
	p-value cluster 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 3	0.657	0.802	0.788	0.641	0.780	0.570	0.838	0.541	0.525	2.052
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
std todos los sujetos		0.666	0.696	0.720	0.702	0.829	0.725	0.759	0.510	0.651	1.210
SCR	std cluster 1	0.576	0.826	0.966	0.799	0.963	1.070	0.406	0.434	0.432	0.374
	p-value cluster 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 2	0.767	0.526	0.420	0.362	0.589	1.582	0.813	0.475	0.687	1.347
	p-value cluster 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 3	0.539	0.816	0.610	0.390	0.951	0.525	0.699	0.467	0.284	0.437
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	0.000	0.326	0.000	0.000	0.000
std todos los sujetos		0.679	0.690	0.703	0.674	0.782	0.705	0.729	0.528	0.664	1.220
SCL	std cluster 1	0.617	0.685	0.709	0.867	1.049	1.000	0.715	0.637	0.699	0.549
	p-value cluster 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 2	0.765	0.639	0.626	0.371	0.491	0.464	0.757	0.449	0.806	1.048
	p-value cluster 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 3	0.645	0.758	0.805	0.670	0.719	0.587	0.837	0.539	0.528	2.022
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.165
std todos los sujetos		0.637	0.677	0.677	0.548	0.787	1.247	0.600	0.535	0.477	0.850
SKT	std cluster 1	0.677	0.917	0.944	0.756	0.741	0.646	0.593	0.569	0.676	0.877
	p-value cluster 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 2	0.648	0.696	0.872	0.742	0.529	0.468	0.737	1.028	0.931	0.843
	p-value cluster 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 3	0.566	0.719	0.843	0.554	0.403	0.625	0.729	0.890	1.026	0.945
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
std todos los sujetos		0.689	0.793	0.868	0.668	0.579	0.585	0.720	0.919	0.870	0.908

ANEXO VIII: Prueba de Levene para la validación de agrupación de sujetos en conducción, vuelta montaña

Test de Levene		Tramo 1	Tramo 2	Tramo 3	Tramo 4	Tramo 5	Tramo 6	Tramo 7	Tramo 8	Tramo 9	Tramo 10
HR	std cluster 1	0.440	0.412	0.510	0.732	0.742	1.139	1.119	0.540	1.227	0.954
	p-value cluster 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 2	0.868	1.058	1.338	1.387	0.555	0.703	0.666	0.752	0.658	0.852
	p-value cluster 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 3	0.909	1.190	0.694	0.873	0.990	1.100	0.951	1.077	1.102	1.047
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 4	0.919	1.019	0.865	1.279	0.395	0.807	0.956	0.921	0.913	0.748
	p-value cluster 4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
std todos los sujetos	0.903	1.103	0.811	0.993	0.922	1.044	0.969	0.966	1.073	0.983	
HRV	std cluster 1	0.880	0.677	0.940	0.842	1.002	1.015	0.959	1.018	1.042	1.018
	p-value cluster 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 2	1.095	1.065	1.235	1.108	0.543	0.623	0.582	0.896	1.045	1.081
	p-value cluster 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 3	0.930	0.965	0.970	1.005	0.994	1.031	1.028	1.053	1.004	1.022
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 4	1.056	0.995	1.002	1.052	0.986	1.021	0.978	0.969	1.010	1.015
	p-value cluster 4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.068	0.000
std todos los sujetos	0.966	0.948	1.033	1.013	0.959	0.994	0.976	1.021	1.022	1.030	
EDA	std cluster 1	0.820	0.787	1.039	0.804	0.841	0.897	1.023	0.825	0.872	0.669
	p-value cluster 1	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 2	0.755	0.808	0.473	0.341	0.522	0.587	0.402	0.346	0.306	0.282
	p-value cluster 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 3	0.793	0.618	0.648	0.495	0.395	0.590	0.705	0.621	0.567	0.654
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 4	1.360	1.001	0.368	0.806	0.854	0.584	0.564	0.734	0.611	0.659
	p-value cluster 4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
std todos los sujetos	1.036	1.049	0.931	0.745	0.751	0.758	0.837	0.739	0.672	0.643	
SCR	std cluster 1	0.782	0.563	0.857	0.726	0.617	0.851	0.843	0.520	1.067	0.669
	p-value cluster 1	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000
	std cluster 2	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.014
	p-value cluster 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 3	0.798	0.817	0.538	0.695	0.645	0.585	0.493	0.584	0.602	0.746
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.990	0.000	0.000
	std cluster 4	1.551	0.756	0.708	0.456	0.404	0.706	0.621	1.082	0.694	0.695
	p-value cluster 4	0.000	0.000	0.000	0.000	0.000	0.000	0.446	0.000	0.000	0.000
std todos los sujetos	0.878	0.703	0.634	0.640	0.572	0.652	0.584	0.641	0.714	0.677	
SCL	std cluster 1	0.566	0.714	0.716	0.636	0.826	0.800	0.936	0.888	0.868	0.407
	p-value cluster 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 2	0.026	0.371	0.547	0.565	0.553	0.503	0.253	0.394	0.284	0.312
	p-value cluster 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 3	0.732	0.607	0.572	0.504	0.389	0.572	0.693	0.576	0.496	0.619
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 4	1.086	0.847	0.271	0.694	0.767	0.345	0.493	0.656	0.409	0.313
	p-value cluster 4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
std todos los sujetos	0.738	0.689	0.693	0.606	0.627	0.649	0.741	0.677	0.588	0.522	
SKT	std cluster 1	1.046	1.429	1.565	0.776	0.332	0.718	0.727	0.444	0.452	0.826
	p-value cluster 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 2	0.730	1.299	1.032	0.731	0.289	0.681	0.675	0.422	0.645	0.689
	p-value cluster 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 3	0.897	1.137	1.230	0.884	0.514	0.545	0.401	0.362	0.555	0.644
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	8.000	0.000	0.000	0.000	0.000
	std cluster 4	1.635	0.636	0.408	0.380	0.288	0.408	0.566	0.512	0.150	0.132
	p-value cluster 4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
std todos los sujetos	1.043	1.204	1.267	0.833	0.618	0.714	0.578	0.500	0.573	0.759	

ANEXO IX: Prueba de Levene para la validación de agrupación de sujetos en conducción, vuelta ciudad

Test de Levene		Tramo 1	Tramo 2	Tramo 3	Tramo 4	Tramo 5	Tramo 6	Tramo 7	Tramo 8	Tramo 9	Tramo 10
HR	std cluster 1	0.963	1.031	1.065	0.788	0.703	0.887	0.774	1.048	1.157	1.130
	p-value cluster 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 2	1.052	0.967	0.775	0.722	0.896	1.101	1.089	1.037	0.998	1.053
	p-value cluster 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 3	1.082	1.058	0.606	0.366	0.602	0.849	0.939	1.261	1.445	0.972
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
std todos los sujetos		1.096	1.034	1.060	0.941	0.886	0.962	0.923	0.890	0.822	0.898
HRV	std cluster 1	1.027	1.079	1.051	1.092	1.070	0.983	0.967	0.813	0.768	0.956
	p-value cluster 1	0.000	0.000	0.119	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 2	1.029	1.035	1.048	1.011	0.931	0.946	0.998	1.033	0.936	0.987
	p-value cluster 2	0.000	0.057	0.000	0.000	0.000	0.000	0.016	0.000	0.012	0.000
	std cluster 3	0.752	0.824	0.930	0.635	0.749	1.121	1.129	1.061	1.024	1.159
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
std todos los sujetos		0.993	0.973	0.975	0.996	1.027	1.035	1.011	1.008	1.016	1.076
EDA	std cluster 1	0.572	0.999	0.736	0.649	0.604	0.555	0.616	0.605	0.538	0.569
	p-value cluster 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 2	0.643	0.701	0.824	0.747	0.719	0.818	0.786	0.748	0.834	0.830
	p-value cluster 2	0.000	0.000	0.000	0.961	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 3	0.465	0.314	0.333	0.285	0.422	0.470	0.721	1.194	0.279	0.595
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
std todos los sujetos		0.666	0.696	0.720	0.702	0.829	0.725	0.759	0.510	0.651	1.210
SCR	std cluster 1	0.725	1.057	0.846	0.619	0.893	0.779	0.711	1.222	1.552	1.235
	p-value cluster 1	0.000	0.000	0.000	0.000	0.451	0.000	0.000	0.000	0.000	0.000
	std cluster 2	0.580	0.608	0.836	0.787	0.901	0.874	0.896	0.707	0.735	1.317
	p-value cluster 2	0.000	0.000	0.536	0.000	0.000	0.134	0.000	0.000	0.000	0.000
	std cluster 3	0.485	0.397	0.365	0.338	0.393	0.462	1.756	1.987	0.428	0.865
	p-value cluster 3	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
std todos los sujetos		0.679	0.690	0.703	0.674	0.782	0.705	0.729	0.528	0.664	1.220
SCL	std cluster 1	0.478	0.780	0.507	0.500	0.452	0.496	0.598	0.673	0.594	0.621
	p-value cluster 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 2	0.476	0.566	0.644	0.549	0.585	0.697	0.663	0.592	0.688	0.681
	p-value cluster 2	0.000	0.000	0.000	0.022	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 3	0.396	0.291	0.255	0.221	0.327	0.651	1.290	1.124	0.200	0.474
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
std todos los sujetos		0.637	0.677	0.677	0.548	0.787	1.247	0.600	0.535	0.477	0.850
SKT	std cluster 1	0.618	0.648	0.800	0.800	0.995	0.746	0.500	0.344	0.611	0.609
	p-value cluster 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	std cluster 2	0.920	0.715	0.730	0.679	0.686	0.689	0.514	0.469	0.481	0.600
	p-value cluster 2	0.000	0.000	0.005	0.000	0.000	0.000	0.009	0.000	0.000	0.000
	std cluster 3	0.151	0.171	0.128	0.233	0.368	0.235	0.308	0.089	0.540	1.212
	p-value cluster 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
std todos los sujetos		0.689	0.793	0.868	0.668	0.579	0.585	0.720	0.919	0.870	0.908

ANEXO X: Prueba F para la validación de agrupación de sujetos en conducción, ida ciudad

Prueba F		Tramo 1	Tramo 2	Tramo 3	Tramo 4	Tramo 5	Tramo 6	Tramo 7	Tramo 8	Tramo 9	Tramo 10
HR	mean cluster 1	-0.385	-0.084	0.233	-0.080	0.265	0.228	0.463	0.177	0.164	0.332
	mean cluster 2	-0.390	0.346	0.150	0.408	0.060	0.214	0.004	-0.186	0.423	0.389
	mean cluster 3	0.088	0.780	-0.074	-0.276	0.388	0.123	-0.093	0.135	-0.071	-0.648
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HRV	mean cluster 1	-0.122	-0.079	-0.163	-0.180	-0.231	-0.107	-0.078	0.080	0.174	0.099
	mean cluster 2	-0.119	-0.214	-0.193	-0.073	-0.076	-0.079	-0.085	-0.037	-0.097	-0.077
	mean cluster 3	0.082	0.365	0.417	0.012	0.277	0.088	-0.113	-0.009	0.082	-0.177
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.009	0.000	0.000
EDA	mean cluster 1	0.799	0.501	0.244	0.928	0.416	0.085	-0.036	0.135	0.158	0.190
	mean cluster 2	-0.299	-0.373	-0.303	-0.129	0.321	0.603	0.545	0.739	0.847	0.381
	mean cluster 3	0.886	0.795	0.844	0.812	0.516	0.404	0.179	0.298	0.267	0.112
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SCR	mean cluster 1	0.840	0.378	0.155	0.438	-0.086	-0.250	-0.159	-0.123	-0.130	-0.025
	mean cluster 2	0.047	0.055	0.113	-0.039	0.115	0.478	0.155	0.347	0.025	-0.075
	mean cluster 3	0.968	0.365	0.257	0.156	0.265	0.063	0.027	0.116	0.562	0.682
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SCL	mean cluster 1	0.562	0.418	0.178	0.805	0.378	0.078	-0.059	0.150	0.119	0.145
	mean cluster 2	-0.441	-0.413	-0.378	-0.199	0.212	0.377	0.435	0.714	0.671	0.272
	mean cluster 3	0.721	0.758	0.787	0.835	0.445	0.581	0.292	0.338	0.174	-0.208
	p-value	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000
SKT	mean cluster 1	0.221	0.221	-0.009	0.235	0.411	0.291	0.319	0.183	0.197	0.201
	mean cluster 2	0.681	0.184	-0.269	-0.207	0.280	0.495	0.456	0.391	0.513	0.451
	mean cluster 3	1.460	0.766	0.629	0.859	0.683	0.219	-0.073	0.155	0.014	-0.053
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

ANEXO XI: Prueba F para la validación de agrupación de sujetos en conducción, ida montaña

Prueba F		Tramo 1	Tramo 2	Tramo 3	Tramo 4	Tramo 5	Tramo 6	Tramo 7	Tramo 8	Tramo 9	Tramo 10
HR	mean clus	0.614	0.242	-0.201	-0.116	0.011	-0.386	-0.175	-0.345	-0.018	0.012
	mean clus	-0.614	0.106	0.107	-0.380	-0.455	0.044	0.045	0.160	0.392	0.311
	mean clus	-0.110	-0.038	0.197	0.071	0.202	0.020	-0.303	0.008	0.274	0.553
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HRV	mean clus	-0.018	0.045	0.146	0.137	0.227	0.328	0.135	0.073	0.103	0.115
	mean clus	-0.080	-0.059	-0.013	0.075	0.130	0.082	0.060	0.087	0.104	0.205
	mean clus	0.310	0.152	0.253	0.222	0.171	0.164	0.079	-0.026	-0.105	-0.010
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
EDA	mean clus	0.212	0.166	0.280	-0.068	-0.048	-0.295	-0.469	-0.612	-0.755	-0.451
	mean clus	0.147	-0.049	-0.175	-0.358	-0.394	-0.392	-0.046	-0.607	-0.425	-0.254
	mean clus	0.225	0.332	0.199	-0.100	-0.106	-0.328	-0.141	-0.522	-0.817	-0.582
	p-value	0.006	0.000	0.000	0.000	0.000	0.143	0.000	0.032	0.000	0.000
SCR	mean clus	-0.026	0.090	0.183	0.036	-0.008	-0.109	-0.364	-0.382	-0.434	-0.296
	mean clus	0.228	0.078	-0.151	-0.238	-0.188	-0.024	-0.053	-0.330	-0.176	0.166
	mean clus	-0.048	0.129	-0.058	-0.186	0.111	-0.114	-0.001	-0.311	-0.367	-0.218
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SCL	mean clus	0.225	0.144	0.223	-0.097	-0.050	-0.300	-0.481	-0.618	-0.774	-0.455
	mean clus	-0.022	-0.171	-0.148	-0.321	-0.277	-0.311	0.061	-0.472	-0.247	-0.055
	mean clus	0.287	0.334	0.236	-0.045	-0.139	-0.332	-0.139	-0.549	-0.838	-0.677
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SKT	mean clus	0.005	0.137	0.050	-0.006	0.028	-0.090	-0.380	-0.492	-0.518	-0.823
	mean clus	0.376	0.487	0.343	0.352	0.221	0.009	-0.125	-0.748	-1.138	-1.470
	mean clus	0.265	0.395	0.491	0.217	0.247	0.243	-0.078	0.024	-0.210	-0.638
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

ANEXO XII: Prueba F para la validación de agrupación de sujetos en conducción, vuelta ciudad

Prueba F		Tramo 1	Tramo 2	Tramo 3	Tramo 4	Tramo 5	Tramo 6	Tramo 7	Tramo 8	Tramo 9	Tramo 10
HR	mean cluster 1	-0.108	0.119	0.224	0.085	-0.136	-0.110	-0.388	-0.017	0.016	-0.114
	mean cluster 2	-0.166	-0.109	-0.463	-0.339	0.012	0.035	0.646	0.165	0.336	0.493
	mean cluster 3	-0.762	-0.351	0.168	-0.081	-0.413	-0.482	-0.539	0.423	-0.213	-0.340
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HRV	mean cluster 1	0.211	0.254	0.173	0.277	0.156	0.023	-0.022	-0.313	-0.322	-0.083
	mean cluster 2	0.264	0.185	0.088	0.055	-0.125	-0.095	-0.009	0.042	-0.153	-0.005
	mean cluster 3	-0.237	-0.132	-0.116	-0.327	-0.262	0.108	0.451	0.663	1.008	1.014
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
EDA	mean cluster 1	0.056	0.228	0.090	0.032	0.178	0.327	0.377	0.556	0.419	0.484
	mean cluster 2	0.006	0.078	0.115	0.128	0.235	0.174	0.156	0.078	0.118	0.292
	mean cluster 3	0.593	0.911	0.773	0.017	0.158	-0.092	1.319	0.638	0.157	0.767
	p-value	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000
SCR	mean cluster 1	-0.007	0.174	0.060	-0.026	0.155	0.072	0.085	0.430	0.464	0.409
	mean cluster 2	-0.067	-0.014	0.092	0.079	0.126	0.092	0.154	-0.048	0.064	0.206
	mean cluster 3	-0.251	-0.131	-0.205	-0.323	-0.263	-0.127	2.146	1.669	0.013	0.075
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SCL	mean cluster 1	0.154	0.351	0.195	0.044	0.142	0.303	0.420	0.513	0.423	0.347
	mean cluster 2	-0.053	0.003	0.007	0.062	0.185	0.188	0.185	0.177	0.196	0.289
	mean cluster 3	0.848	1.125	1.020	0.244	0.360	-0.057	0.024	-0.337	0.163	0.830
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SKT	mean cluster 1	0.483	0.751	0.522	0.388	0.311	0.213	0.615	0.555	0.293	0.651
	mean cluster 2	0.461	0.624	0.659	0.797	0.613	0.539	0.432	0.209	0.157	0.014
	mean cluster 3	-0.129	0.487	1.005	1.004	1.300	1.298	0.533	0.057	0.321	0.102
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

ANEXO XIII: Prueba F para la validación de agrupación de sujetos en conducción, vuelta montaña

Prueba F		Tramo 1	Tramo 2	Tramo 3	Tramo 4	Tramo 5	Tramo 6	Tramo 7	Tramo 8	Tramo 9	Tramo 10
HR	mean cluster 1	-0.418	-0.415	-0.338	0.063	0.141	-0.097	-0.095	-0.341	0.194	0.002
	mean cluster 2	0.679	0.529	0.015	0.254	0.654	0.368	-0.260	0.578	0.657	-0.297
	mean cluster 3	0.100	0.358	-0.069	-0.113	-0.158	0.025	-0.192	0.129	0.265	0.088
	mean cluster 4	0.626	0.837	-0.408	0.240	-0.799	0.013	0.227	0.143	0.205	-0.151
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.008
HRV	mean cluster 1	-0.217	-0.486	-0.183	-0.405	-0.109	0.066	0.109	0.245	0.281	0.312
	mean cluster 2	-0.042	0.058	0.504	0.054	-0.414	-0.356	-0.406	-0.054	0.316	0.360
	mean cluster 3	-0.181	-0.103	-0.102	-0.008	-0.021	0.090	0.134	0.184	0.137	0.196
	mean cluster 4	0.207	0.033	0.257	0.264	0.142	0.180	0.101	0.204	0.291	0.257
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
EDA	mean cluster 1	-0.515	-0.656	-0.394	-0.260	-0.795	-0.148	-0.327	-0.310	-0.262	0.043
	mean cluster 2	0.904	1.490	1.105	0.794	0.759	0.439	0.409	0.338	0.108	-0.262
	mean cluster 3	-0.788	-0.774	-0.729	-0.514	-0.427	-0.543	-0.514	-0.345	-0.146	0.046
	mean cluster 4	-0.274	-0.456	-0.147	-0.640	-0.210	0.143	0.436	0.571	0.567	0.414
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SCR	mean cluster 1	0.112	-0.223	0.015	-0.090	-0.204	0.142	-0.156	-0.104	0.066	0.133
	mean cluster 2	-0.120	-0.120	-0.120	-0.120	-0.120	-0.120	-0.120	-0.120	-0.120	-0.120
	mean cluster 3	-0.249	-0.218	-0.377	-0.220	-0.254	-0.280	-0.353	-0.279	-0.212	0.005
	mean cluster 4	0.106	-0.147	-0.021	-0.456	-0.213	-0.066	-0.077	0.352	0.058	0.125
	p-value	0.000	0.045	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000
SCL	mean cluster 1	-0.610	-0.584	-0.682	-0.324	-0.762	-0.131	-0.316	-0.225	-0.246	-0.018
	mean cluster 2	-0.121	0.026	0.365	0.137	0.060	0.018	-0.162	-0.018	-0.177	-0.144
	mean cluster 3	-0.824	-0.772	-0.736	-0.497	-0.372	-0.503	-0.470	-0.335	-0.167	-0.014
	mean cluster 4	-0.779	-0.638	-0.206	-0.563	-0.253	0.111	0.351	0.286	0.265	0.143
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SKT	mean cluster 1	-0.027	-0.480	-0.714	-1.357	-1.173	-1.198	-0.499	-0.319	0.098	0.447
	mean cluster 2	-0.618	-1.040	-1.632	-0.743	0.287	0.182	-0.123	0.043	0.006	0.966
	mean cluster 3	-0.572	-0.582	-0.706	-0.795	-0.884	-0.764	-0.703	-0.736	-0.378	-0.095
	mean cluster 4	-0.414	-0.997	-1.039	-0.998	-1.005	-0.739	-0.501	-0.230	0.216	0.326
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

ANEXO XIV: Interpretacion según niveles de significancia

Nivel de significancia	Interpretación
0.1	marginalmente significativo
0.05	significativo
0.01	muy significativo

ANEXO XV: Tabla de contingencia de fatiga SNA vs grupos para cada subactividad

a) Viaje de ida, trecho ciudad

	Grupo 1	Grupo 2	Grupo 3	Total
Awake	2472	2364	1919	6755
Stressed	479	480	257	1216
Fatigued	187	153	268	608
Drowsy	456	484	274	1214
Total	3594	3481	2718	9793

b) Viaje de ida, trecho montaña

	Grupo 1	Grupo 2	Grupo 3	Total
Awake	2538	3697	1652	7887
Stressed	520	675	319	1514
Fatigued	202	277	122	601
Drowsy	620	760	358	1738
Total	3880	5409	2451	11740

c) Viaje de vuelta, trecho montaña

	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Total
Awake	1829	967	4339	1126	8261
Stressed	246	128	757	279	1410
Fatigued	163	104	324	99	690
Drowsy	228	116	783	315	1442
Total	2466	1315	6203	1819	11803

d) Viaje de vuelta, trecho ciudad

	Grupo 1	Grupo 2	Grupo 3	Total
Awake	1773	4315	677	6765
Stressed	322	771	64	1157
Fatigued	143	335	104	582
Drowsy	300	798	79	1177
Total	2538	6219	924	9681

ANEXO XVI: Tabla de contingencia de fatiga de rendimiento vs grupos para cada subactiviad

a) Viaje de ida, trecho ciudad

	Grupo 1	Grupo 2	Grupo 3	Total
Sin fatiga	2075	1771	1300	5146
Baja fatiga	1044	1339	1049	3432
Alta fatiga	475	371	369	1215
Total	3594	3481	2718	9793

b) Viaje de ida, trecho montaña

	Grupo 1	Grupo 2	Grupo 3	Total
Sin fatiga	466	229	291	986
Baja fatiga	354	138	267	759
Alta fatiga	113	25	64	202
Total	933	392	622	1947

c) Viaje de vuelta, trecho montaña

	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Total
Sin fatiga	1473	835	3643	1079	7030
Baja fatiga	739	336	1809	509	3393
Alta fatiga	254	144	751	231	1380
Total	2466	1315	6203	1819	11803

d) Viaje de vuelta, trecho ciudad

	Grupo 1	Grupo 2	Grupo 3	Total
Sin fatiga	1393	3050	463	4906
Baja fatiga	709	2446	333	3488
Alta fatiga	436	723	128	1287
Total	2538	6219	924	9681

ANEXO XVII: Prueba de Levene, Fatiga SNA

Característica	Levene Fatiga SNA							
	IDA CIUDAD		IDA MONTAÑA		VUELTA CIUDAD		VUELTA MONTAÑA	
	Tendencia	p-valor	Tendencia	p-valor	Tendencia	p-valor	Tendencia	p-valor
meanHR		0.000		0.324		0.000		0.000
stdHR		0.009		0.016		0.133		0.000
rangeHR		0.074		0.067		0.727		0.000
meanHRV		0.001		0.000		0.000		0.000
stdHRV		0.000		0.000		0.000		0.000
rmssdHRV		0.000		0.000		0.000		0.000
VLF		0.000		0.000		0.000		0.000
TP		0.000		0.000		0.000		0.000
ratioLFHF		0.000		0.000		0.000		0.000
meanEDA		0.000		0.000		0.000		0.001
stdEDA		0.006		0.039		0.000		0.000
meanSCL		0.000		0.000		0.007		0.000
meanSCR		0.000		0.055		0.000		0.000
rangeSCL		0.001		0.919		0.000		0.538
maxSCR		0.000		0.868		0.005		0.000
meanSKT		0.000		0.000		0.000		0.001
stdSKT		0.001		0.195		0.000		0.000
rangeSKT		0.050		0.050		0.001		0.000
Mean_humedad		0.024		0.002		0.035		0.000
std_humedad		0.620		0.000		0.002		0.000
Mean_temp_DFT		0.034		0.000		0.039		0.115
std_temp_DFT		0.452		0.000		0.009		0.000
Mean_temp_CSS		0.022		0.664		0.582		0.000
std_temp_CSS		0.000		0.000		0.000		0.000
Mean_co2		0.000		0.400		0.001		0.906
std_co2		0.000		0.111		0.000		0.008
Mean_sonido		0.000		0.320		0.000		0.000
std_sonido		0.000		0.001		0.991		0.000
mean_acc		0.055		0.000		0.054		0.000
Min_speed		0.020		0.180		0.722		0.598
Max_speed		0.000		0.001		0.000		0.000
Mean_speed		0.000		0.021		0.000		0.000
std_speed		0.025		0.816		0.000		0.002
Fe		0.052		0.006		0.391		0.018
Ne		0.000		0.000		0.000		0.113

ANEXO XVIII: Prueba de Levene, Fatiga rendimiento

Levene Fatiga rendimiento

Características	IDA CIUDAD		IDA MONTAÑA		VUELTA CIUDAD		VUELTA MONTAÑA	
	Tendencia	p-valor	Tendencia	p-valor	Tendencia	p-valor	Tendencia	p-valor
meanHR		0.000		0.000		0.005		0.000
stdHR		0.000		0.003		0.217		0.000
rangeHR		0.000		0.002		0.294		0.000
meanHRV		0.001		0.000		0.536		0.000
stdHRV		0.000		0.000		0.000		0.000
rmsdHRV		0.000		0.000		0.000		0.021
VLF		0.000		0.358		0.000		0.000
TP		0.000		0.432		0.000		0.000
ratioLFHF		0.641		0.001		0.370		0.673
LF		0.000		0.457		0.000		0.000
HF		0.000		0.000		0.000		0.000
meanEDA		0.000		0.000		0.511		0.162
stdEDA		0.000		0.023		0.000		0.000
meanSCL		0.000		0.000		0.000		0.001
meanSCR		0.022		0.000		0.000		0.000
rangeSCL		0.000		0.000		0.000		0.000
maxSCR		0.001		0.000		0.000		0.000
meanSKT		0.000		0.000		0.000		0.003
stdSKT		0.000		0.000		0.088		0.000
rangeSKT		0.000		0.000		0.068		0.000
Mean_humedad		0.000		0.000		0.000		0.000
std_humedad		0.000		0.000		0.000		0.001
Mean_temp_DFT		0.000		0.000		0.000		0.000
std_temp_DFT		0.001		0.000		0.000		0.000
Mean_temp_CSS		0.000		0.074		0.000		0.000
std_temp_CSS		0.000		0.000		0.010		0.006
Mean_co2		0.000		0.071		0.000		0.000
std_co2		0.000		0.000		0.751		0.000
Mean_sonido		0.000		0.000		0.000		0.000
std_sonido		0.000		0.000		0.000		0.000
mean_acc		0.000		0.000		0.000		0.000
Min_speed		0.000		0.000		0.000		0.000
Max_speed		0.000		0.000		0.000		0.000
Mean_speed		0.000		0.000		0.000		0.000
std_speed		0.000		0.000		0.111		0.000

ANEXO XIX: Ejemplo de interpretación de la tendencia para la Fatiga SNA y de rendimiento

