



INSTITUTO MEXICANO DEL TRANSPORTE

Desarrollo de un software para la generación de modelos de evolución del IRI basados en cadenas de Márkov

José Ricardo Solorio Murillo
Etzael Contreras Jácome
Montserrat Montoya Ortega

Publicación Técnica No. 625
San Fandila, Qro.
2021

ISSN 0188-7297

Esta investigación fue realizada en la Coordinación de Infraestructura de Vías Terrestres del Instituto Mexicano del Transporte por el Dr. José Ricardo Solorio Murillo, el M. I. Etzael Contreras Jácome y Monserrat Montoya Ortega.

La presente publicación es el producto final del proyecto de investigación interna II 01/17, “Desarrollo de un software para la generación de modelos de evolución del IRI basados en cadenas de Márkov, fase 2”

Como parte de este proyecto, el M. I. Contreras desarrolló la tesis de maestría titulada “Desarrollo de un software para el análisis estadístico de series históricas del IRI”, la cual le permitió obtener el grado de Maestro en Vías Terrestres por la Facultad de Ingeniería Civil de la Universidad Veracruzana, Región Xalapa.

Contenido

	Página
Índice de figuras	v
Índice de tablas	vii
Sinopsis.....	ix
Abstract	xi
Resumen ejecutivo.....	xiii
Introducción.....	1
1 Antecedentes	3
2 Análisis estadístico.....	15
3 Desarrollo del software.....	25
4 Comparación de las pruebas	51
5 Conclusiones.....	59
Bibliografía	61

Índice de figuras

Figura 1.1 Modelo de cuarto de carro (Romero & Lozano, 1995).	5
Figura 1.2 Escala de Valores del IRI y características de los pavimentos.	6
Figura 1.3 Prevalencia de la hipótesis nula en presencia de obra.	11
Figura 1.4 Matrices de probabilidades de transición.	13
Figura 2.1 Ejemplo de la asimetría de las distribuciones de diferencias de IRI.	20
Figura 3.1 Diagrama de caja de los datos de la familia <i>00-ECa-DCm-PMa</i>	31
Figura 3.2 Evolución del valor esperado del IRI, familia <i>00-ECa-DCm-PMa</i>	33
Figura 3.3 Diagrama de flujo de los procesos del software.	34
Figura 3.4 Vista parcial del código del programa en la plataforma de VBA.	36
Figura 3.5 Fragmento de código relacionado con el algoritmo GRG.	41
Figura 3.6 Modelos de la familia <i>00-ECa-DCm-PMa</i> , prueba de los signos.	46
Figura 3.7 Modelos de la familia <i>00-ECa-DCm-PMa</i> , prueba t por pares.	46
Figura 3.8 Pantalla de bienvenida del software.	48
Figura 3.9 Menú principal.	48
Figura 3.10 Diálogo para la definición de parámetros estadísticos.	49
Figura 4.1 Valor esperado, familia <i>00-ECb-DCb-PMa</i> , prueba de los signos.	53
Figura 4.2 Valor esperado, familia <i>00-ECb-DCb-PMa</i> , prueba t por pares.	54
Figura 4.3 Valor esperado, familia <i>00-ECm-DCa-PMa</i> , prueba de los signos.	55
Figura 4.4 Valor esperado, familia <i>00-ECm-DCa-PMa</i> , prueba t por pares.	55
Figura 4.5 Valor esperado, familia <i>00-ECm-DCm-PMa</i> , prueba de los signos.	57
Figura 4.6 Valor esperado, familia <i>00-ECm-DCm-PMa</i> , prueba t por pares.	57

Índice de tablas

Tabla 1.1 Rangos de clasificación por parámetro de los tramos.	7
Tabla 1.2 Resultado de la clasificación de los tramos en familias y grupos.	9
Tabla 2.1 Elementos de la prueba de los signos.	17
Tabla 2.2 Elementos de la prueba t por pares.	18
Tabla 2.3 Niveles de asimetría en el grupo <i>ECb-DCb-PMm / 1993-2009-2012</i>	21
Tabla 2.4 Posibles resultados de una prueba de hipótesis.	21
Tabla 3.1 Aplicación de la prueba de los signos, familia <i>00-ECa-DCm-PMa</i>	27
Tabla 3.2 Rangos de clasificación del IRI para la generación de las MPT.	28
Tabla 3.3 Matrices de probabilidades de transición, familia <i>00-ECa-DCm-PMa</i> . ..	28
Tabla 3.4 Cadenas de Márkov, familia <i>00-ECa-DCm-PMa</i>	29
Tabla 3.5 Valores esperados de IRI en m/km, familia <i>00-ECa-DCm-PMa</i>	31
Tabla 3.6 Resultados de la prueba de los signos, familia <i>00-ECa-DCm-PMa</i>	39
Tabla 3.7 Resultados de la prueba t por pares, familia <i>00-ECa-DCm-PMa</i>	39
Tabla 3.8 Obtención de las MPT, familia <i>00-ECa-DCm-PMa</i>	42
Tabla 3.9 Desarrollo de las cadenas de Márkov, familia <i>00-ECa-DCm-PMa</i>	43
Tabla 3.10 Cálculo de valores esperados, familia <i>00-ECa-DCm-PMa</i>	45
Tabla 4.1 Años útiles y errores según el tipo de prueba.	52
Tabla 4.2 Valores P de la familia <i>00-ECb-DCb-PMa</i>	53
Tabla 4.3 Valores P de la familia <i>00-ECm-DCa-PMa</i>	54
Tabla 4.4 Valores P de la familia <i>00-ECm-DCm-PMa</i>	56

Sinopsis

El presente documento describe el proceso de desarrollo de un software para auxiliar en la generación de modelos de evolución del IRI basados en datos históricos y cadenas de Márkov, que emplea una metodología integrada en trabajos previos llevados a cabo en la Coordinación de Infraestructura de Vías Terrestres del IMT. Específicamente, el software automatiza las partes del procedimiento de modelización relativas a la identificación de ciclos de deterioro del pavimento y a la generación de las cadenas de Márkov. Además, el software incorpora la prueba *t por pares* como un mecanismo alternativo a la *prueba de los signos* para identificar las fronteras de los ciclos. Cuando se emplea la prueba de los signos, la herramienta desarrollada produce resultados prácticamente idénticos a los obtenidos manualmente, pero en un tiempo considerablemente menor. Con respecto a la prueba *t por pares*, en general su aplicación se traduce en una mayor capacidad para identificar datos que, en realidad, no son útiles para la modelización del deterioro; lo cual se relaciona con la mayor potencia de esta prueba. Aunque, por un lado, los datos no cumplen con el supuesto de normalidad de la prueba *t*, por otro tienen las características mínimas necesarias para producir resultados confiables.

Abstract

This document describes the process of developing a software to help in the generation of IRI evolution models based on historical data and Markov chains, using a methodology integrated in previous work carried out in the Department of Land Transport Infrastructure of IMT. Specifically, the software automates the parts of the modelling procedure related to the identification of pavement deterioration cycles and the generation of Markov chains. In addition, the software incorporates the *paired t-test* as a mechanism alternative to the *sign test* for identifying cycle boundaries. When the sign test is used, the developed tool produces results almost identical to those obtained manually, but in a considerably shorter time. With respect to the paired t-test, in general its application translates into a greater capacity to identify data that are not actually useful for deterioration modeling, which is related to the greater power of this test. Although, on the one hand, data do not comply with the t test requirement requiring a normal distribution, on the other hand they have the minimum characteristics to ensure the reliability of the results.

Resumen ejecutivo

En años previos, el Grupo de Investigación en Gestión de Infraestructura Carretera del Instituto Mexicano del Transporte llevó a cabo una serie de trabajos que culminaron con la integración de una metodología basada en datos históricos y en cadenas de Márkov para la modelización de la evolución temporal del IRI a nivel de red. Esta metodología comprende tres fases principales:

- a) Clasificar los tramos de carretera a los que se refieren los datos en familias y grupos, a partir de las características del pavimento y de los registros de conservación existentes.
- b) Aplicar la *prueba de los signos* para identificar los años en los que podrían haberse realizado trabajos de conservación u ocurrido errores de medición o de procesamiento de datos, y asumir que tales años constituyen los límites de ciclos de deterioro dentro de las series temporales.
- c) Generar cadenas de Márkov para los ciclos de deterioro identificados.

El deterioro de los pavimentos se describe normalmente en términos de tres atributos: capacidad funcional, capacidad estructural y deterioros superficiales. La capacidad funcional se refiere al nivel de servicio ofrecido a los usuarios, que se mide a través de indicadores relacionados con la comodidad y la seguridad del viaje, como el Índice de Regularidad Internacional (IRI), el coeficiente de fricción y la profundidad de la macrotextura. La capacidad estructural evalúa si el pavimento, en su estado actual, es adecuado para soportar las sollicitaciones impuestas por el tránsito y las condiciones climáticas para una periodo de vida útil determinado. Finalmente, a través de los deterioros superficiales, se determina si la presencia y gravedad de alteraciones de la capa de rodadura como agrietamiento, baches o deformaciones transversales compromete de alguna manera la capacidad funcional del pavimento.

En un primer ejercicio de aplicación de la metodología, se eligió el IRI como parámetro central de la modelización porque constituye un índice muy útil para describir la calidad del servicio ofrecido y, en este sentido, ha sido utilizado durante los últimos años por dependencias como la Dirección General de Conservación de Carreteras de la SCT y Caminos y Puentes Federales de Ingresos y Servicios Conexos.

Como se mencionó previamente, la metodología desarrollada en el IMT propone modelizar la evolución del IRI a través del enfoque de cadenas de Márkov. Este enfoque se refiere a un proceso estocástico que describe el estado de un sistema en el tiempo, y que tiene las siguientes propiedades: i) El tiempo se representa como

una variable discreta que define ciclos de servicio con una duración dada que, en el caso de los pavimentos, es normalmente de un año; ii) El proceso consiste en un espacio finito de estados posibles, que pueden definirse a través de intervalos de clase; iii) El cambio entre estados está determinado por un conjunto de probabilidades denominadas probabilidades de transición; iv) La probabilidad de que el sistema pase del estado i al estado j durante un ciclo de servicio, depende únicamente de su estado actual y no de los estados anteriores.

La modelización mediante cadenas de Márkov se basa en dos entidades fundamentales: el *vector de estado* y la *matriz de probabilidades de transición (MPT)*. La primera de ellas define una distribución de las probabilidades de que los tramos se encuentren en alguno de los estados posibles. La MPT concentra todas las probabilidades de que un tramo permanezca en su estado actual o cambie a otro más desfavorable. Una cadena de Márkov es la serie de vectores de estado que describen la evolución de un determinado indicador de la condición del pavimento en el tiempo. Tanto los vectores de estado como las matrices de probabilidades de transición se calculan a partir de información histórica. Las MPT pueden obtenerse con alguno de los siguientes tres métodos: proporciones de segmentos que experimentan determinados cambios de estado, minimización del error en el cálculo de los vectores de estado y minimización del error al estimar los valores esperados de IRI.

La información histórica necesaria para la modelización del IRI consiste en series de datos que describen ciclos de deterioro, es decir, periodos en los que este índice se incrementa. Las fronteras de los ciclos de deterioro corresponden a los años en los que se hayan efectuado acciones de construcción nueva, conservación periódica o rehabilitación, aunque también pueden estar determinados por series anuales de datos con errores de medición que produzcan una reducción artificial del IRI.

Para identificar los límites de los ciclos de deterioro, la metodología propone la aplicación de una prueba de hipótesis basada en la prueba de los signos, en la que la hipótesis nula describe el caso en el que la diferencia entre mediciones consecutivas de IRI es igual a cero y la hipótesis alternativa indica que tal diferencia es negativa, es decir, que existe evidencia estadísticamente significativa de que se interrumpió el proceso de degradación del pavimento. Cuando esto ocurre, se cierra el ciclo de deterioro y se descarta el resto de los datos para efectos de la definición de cada ciclo.

Originalmente, se había considerado utilizar la *prueba t por pares* para la delimitación de los ciclos de deterioro, ya que se trata de una de las pruebas más efectivas para evaluar hipótesis relacionadas con la diferencia entre las medias de dos poblaciones. Sin embargo, esta prueba requiere que la diferencia entre los pares de mediciones se distribuya normalmente, condición que no se cumple, en la mayoría de los casos, para la información disponible. Por esta razón, se decidió en primera instancia sustituir la prueba t por la prueba de los signos, una prueba no paramétrica que no tiene ningún requisito relacionado con la distribución de los

datos, aunque tiene la desventaja de que su potencia es menor que la de la prueba t por pares.

La potencia de una prueba de hipótesis se refiere a la probabilidad de rechazar la hipótesis nula cuando es falsa, lo que, en el caso del problema abordado en este trabajo, equivale a detectar el límite de un ciclo de deterioro en el año en el que efectivamente ocurre. Lo contrario da lugar a un error tipo II, y conlleva la incorporación a un ciclo de deterioro de datos que deberían ser excluidos.

Como parte del trabajo que se presenta en este documento, se hizo una investigación bibliográfica complementaria para evaluar en forma más detallada la posibilidad de aplicar la prueba t por pares. A este respecto, se encontró que la prueba funciona adecuadamente, aun violando el supuesto de normalidad, cuando la distribución subyacente es simétrica, unimodal y continua. En lo que respecta a la primera propiedad, pudo comprobarse que las diferencias entre las mediciones consecutivas tienen un nivel de asimetría de bajo a moderado en la mayoría de los casos. Por otro lado, el carácter unimodal de la distribución de las diferencias se garantiza en gran medida con la fase inicial de clasificación de los tramos. Finalmente, es claro que la distribución de las diferencias puede considerarse una función continua.

El software desarrollado automatiza las siguientes funciones previstas en la metodología de modelización:

- a) Identificación de los ciclos de deterioro o de la porción de datos útiles de las series disponibles para la generación de los modelos, ofreciendo para ello al usuario la opción de utilizar la prueba de los signos que se propuso originalmente o la prueba t por pares.
- b) Obtención de los modelos de evolución del IRI basados en cadenas de Márkov, con base en los tres métodos ya mencionados para la conformación de las MPT: proporciones de segmentos que experimentan determinados cambios de estado, minimización del error en el cálculo de los vectores de estado y minimización del error al estimar los valores esperados de IRI.
- c) Representación gráfica de los valores esperados de IRI.

Debido al formato de los datos originales, se omitió el desarrollo de una componente para la clasificación inicial de los tramos en familias y grupos. Sin embargo, es probable que esta componente se retome en el futuro.

La programación de las funciones antes enumeradas se realizó en el lenguaje *Visual Basic for Applications (VBA)* de Excel. Entre los factores que condujeron al uso de esta herramienta pueden mencionarse los siguientes:

- Tanto los datos de entrada como una parte importante de los resultados del proceso de generación de los modelos tienen un formato tabular, propicio para su procesamiento en Excel.

- Esta plataforma proporciona una rica biblioteca de funciones matemáticas y estadísticas que pueden llamarse desde los programas escritos en VBA. Las funciones disponibles comprenden las del cálculo de los parámetros de estadística descriptiva y otras variables necesarias para la determinación del *valor P* de pruebas de hipótesis. Adicionalmente, desde el código de VBA de Excel puede utilizarse el *algoritmo de optimización no lineal del gradiente reducido generalizado (GRG)*, disponible en el paquete de herramientas *Solver* de Excel.
- Otro recurso muy importante al que se puede recurrir en el desarrollo de programas de VBA de Excel son las funciones de la biblioteca de gráficos, lo que simplifica notablemente la preparación de las salidas visuales de los programas.

Tomando en cuenta lo anterior, se codificaron e integraron las funciones de identificación de ciclos de deterioro, generación de las cadenas de Márkov y comparación gráfica de los valores esperados de IRI, junto con una interfaz amigable para el uso del software. La herramienta desarrollada se aplicó para generar modelos con los mismos datos utilizados en la formulación de la metodología, a fin de verificar que los cálculos automatizados se realizaran en forma correcta, por lo menos en lo que toca a la prueba de los signos.

En relación con los diferentes resultados producidos, se destaca lo siguiente:

- a) En general, los modelos que se obtienen cuando se aplica la prueba de los signos con el software son prácticamente idénticos a los que se generaron en forma manual durante el desarrollo de la metodología.
- b) El tiempo de obtención de los modelos se reduce de días a minutos.
- c) Al aplicar la prueba t por pares, se observan reducciones en la longitud de algunos ciclos de deterioro o, dicho en otras palabras, un aumento en el rechazo de series anuales durante la integración de los ciclos. Este hecho, está relacionado con la mayor potencia de la prueba, la cual reduce la probabilidad de ocurrencia de los errores tipo II.
- d) Relacionado con lo anterior, al optar por la prueba t, en la mayoría de los casos se observa una disminución muy importante del error en la estimación de los modelos, la cual alcanza niveles de hasta 70 %.

Como conclusión general, el software desarrollado constituye una contribución significativa para propiciar la aplicación de técnicas estadísticas en la modelización del IRI a nivel de red para las carreteras mexicanas.

Introducción

La evaluación del estado de la red federal de carreteras es una tarea esencial para identificar las acciones de conservación o rehabilitación necesarias en los diferentes tramos. Esta evaluación implica la medición de parámetros diversos como el Índice de Regularidad Internacional (IRI), los deterioros superficiales y las deflexiones del pavimento. Así mismo, conlleva el proceso de esta información para diagnosticar la situación actual y pronosticar la situación futura.

A finales de 2014, la red nacional de caminos tenía una longitud total de 389,345 km, de los que 134,305 correspondían a carreteras libres, 9,457 a carreteras de cuota, 175,775 a caminos rurales y alimentadores y 69,808 a brechas mejoradas (Martínez, et al., 2017). Asimismo, se estimaba que, para 2019, aproximadamente el 79 % del transporte se realizaría por estas carreteras y que, de este, el 83 % correspondería a movimiento de carga y el 97 % a pasajeros. Estas cifras son indicativas de la importancia de la conservación de la red de carreteras para el desarrollo económico del país

El Índice de Regularidad Internacional (IRI) fue propuesto por el Banco Mundial en 1986 como un estándar para caracterizar la regularidad superficial de los pavimentos, y ha sido utilizado por la Secretaría de Comunicaciones y Transportes (SCT) durante las últimas dos décadas. En el pasado, los indicadores de la regularidad podían ser diferentes según el equipo utilizado para su medición; una de las grandes ventajas del IRI es que puede medirse con equipos diversos sin que ello impida hacer comparaciones entre la regularidad de diferentes carreteras o dar seguimiento a la evolución de este parámetro en el tiempo (Arriaga, et al., 1998).

La norma ASTM E 867 sobre la terminología relacionada con las características de la superficie recorrida (ASTM, 2012), define la irregularidad (*roughness*) como la desviación de la superficie de un pavimento con respecto a una perfectamente plana, con dimensiones que afectan la dinámica del vehículo y la calidad de la conducción. Aunque, con frecuencia, se utiliza en español el término “rugosidad” para referirse a esta propiedad, en el presente documento se empleará como alternativa el vocablo “regularidad” que, aunque se refiere al atributo opuesto, es igualmente útil para describir esta característica de los pavimentos, la cual se relaciona primordialmente con la uniformidad de la superficie de rodadura y no con sus propiedades de resistencia al deslizamiento, como lo podría sugerir el término “rugosidad”.

Desde hace más veinte años, los responsables de la red federal mexicana de carreteras recopilan indicadores del estado de los pavimentos de la red, especialmente el IRI, parámetro que ha sido ampliamente utilizado para caracterizar

de manera general el estado de las carreteras federales y determinar sus necesidades de conservación. Esta información tiene un potencial para modelar la evolución del IRI a nivel de red que hasta hace algunos años no había sido evaluado.

La preparación de programas anuales y plurianuales de conservación de pavimentos requiere que la medición del estado del pavimento se realice en forma periódica. Por esta razón, en años recientes la Dirección General de Servicios Técnicos de la SCT formalizó el proceso de recolección de datos a través un programa específico de evaluación de carreteras que abarca a la red federal libre de peaje y a la de cuota (SCT, 2018). El objetivo de este programa ha sido determinar las condiciones funcionales, estructurales y de seguridad de la red, con base en lo establecido en la Normativa para la Infraestructura del Transporte, a efecto de contar información suficiente para definir las soluciones más convenientes para la modernización y conservación del patrimonio vial a cargo de la SCT.

Junto con la información de campo, la preparación de programas de conservación requiere de modelos matemáticos para pronosticar la evolución del estado de los pavimentos en el futuro y, con ello, anticipar las obras que será necesario realizar para revertir el deterioro debido a la repetición de las cargas del tránsito, a los agentes medioambientales y a otros factores.

La experiencia mexicana en el ámbito de los modelos de deterioro de pavimentos se limita a los de métodos de diseño desarrollados en el país: por un lado, el del Instituto de Ingeniería de la UNAM, implementado a través del programa de cómputo *DISPAV* (Corro, et al., 1981) y, por otro, el del Instituto Mexicano del Transporte, conocido como *IMT-PAVE* por el nombre del software respectivo (Garnica & Hernández, 2013). Sin embargo, no existe experiencia nacional en lo que se refiere a modelos de deterioro a nivel de red, que son los necesarios para la estimación de las necesidades de conservación de pavimentos consideradas en los programas anuales o plurianuales.

Este hecho constituyó la principal motivación para la realización en el IMT de un proyecto para la aplicación de métodos estadísticos en el desarrollo de modelos de deterioro para el nivel de red a partir de información histórica (Solorio, 2018). El trabajo que se presenta en este documento describe el desarrollo de una serie de rutinas en el lenguaje *Visual Basic for Applications* de Excel para la automatización de los procesos de identificación de ciclos de deterioro y de cálculo de cadenas de Márkov que se integraron en el proyecto citado.

1. Antecedentes

1.1 Evaluación de la red federal

En el año 2012 la SCT, a través de Dirección General de Servicios Técnicos, decidió sistematizar los esfuerzos que realiza desde la década de los 90 para obtener datos objetivos del estado de los pavimentos (SCT, 2018). Para ello, lanzó el *Programa de Auscultación de la Red Carretera Federal*.

De acuerdo con la fuente antes citada, este programa fue diseñado con el objeto de determinar las condiciones superficiales, estructurales y de seguridad vial de las carreteras del país, través de la obtención de los siguientes indicadores: Índice de Regularidad Internacional (IRI), profundidad de roderas, deterioros superficiales, profundidad de la macrotextura, coeficiente de fricción y deflexiones del pavimento.

El programa de auscultación se ha llevado a la práctica de manera ininterrumpida desde el año 2013 en los treinta y un estados del país, y se encuentra registrado en la cartera de inversiones de la Secretaría de Hacienda y Crédito Público bajo el rubro *Estudios de preinversión*. Con ello, se asegura la disponibilidad de información en el futuro para avanzar en el estudio y modelización del deterioro de pavimentos en México.

El proyecto que da sustento conceptual al trabajo que se presenta en este documento se inició en 2013, de manera que en él se utilizan series históricas de un periodo anterior al que abarca el programa de auscultación antes mencionado. Aunque, a fin de verificar la correcta implementación de los procedimientos de cálculo, en el desarrollo del software se utilizaron los mismos datos del proyecto referido, es claro que tanto la metodología de análisis como el programa de cómputo podrán utilizarse con la información más reciente.

En el proyecto original se contempló únicamente la modelización del IRI. Lo anterior se debió a dos factores principales: por una parte, el examen de la información disponible permitió concluir que con este parámetro se tenía la mayor cobertura espacial y temporal y, por otra, se trata de un indicador utilizado con frecuencia, a nivel nacional e internacional, para describir el estado de los pavimentos de manera general. Asimismo, el IRI tiene una influencia determinante en los costos de los usuarios de la infraestructura vial, así como en el nivel de confort que experimentan cuando circulan por las carreteras. De esta manera, el IRI constituye también una medida del nivel de servicio proporcionado por la organización responsable.

En el siguiente inciso, se presentan algunos detalles adicionales sobre la naturaleza y la génesis del IRI.

1.2 Índice de Regularidad Internacional (IRI)

El Índice de Regularidad Internacional, conocido en todo el mundo por las siglas de su nombre en inglés (*IRI, International Roughness Index*), fue propuesto por el Banco Mundial en 1986, a partir de trabajos previos auspiciados por el Programa Nacional de Investigación Cooperativa de Carreteras de los EUA (Gillespie, et al., 1980). Desde entonces, este índice se fue consolidando paulatinamente como un parámetro de referencia para caracterizar la regularidad de un camino.

En la década de los setenta, el Banco Mundial patrocinó varios programas de investigación para el análisis técnico-económico de carreteras en países en desarrollo. Algunos de ellos demostraron que los caminos con poca inversión no son rentables para dichas naciones, debido a que los correspondientes costos de usuario son comparativamente altos. También, permitieron identificar a la regularidad del pavimento como un factor determinante del nivel de esos costos. Adicionalmente, se advirtió que los datos de regularidad disponibles no podían ser comparados entre sí, ya que, aún dentro de un mismo país, las mediciones se realizaban con métodos distintos.

En 1982, el Banco Mundial inició un estudio adicional con el propósito de establecer un estándar de calibración para las mediciones de regularidad (Sayers, et al., 1986-2). Durante el estudio, se observó que era posible obtener correlaciones entre los valores de regularidad recolectados con diferentes equipos, lo que dio pie a la búsqueda de un índice de referencia que a la postre condujo al Índice de Regularidad Internacional.

El IRI es un estadístico del perfil de la superficie de rodadura de un pavimento que no depende del equipo utilizado, por lo que puede considerarse una propiedad intrínseca del perfil superficial. Las ecuaciones para su cálculo fueron desarrolladas y evaluadas bajo la premisa de minimizar los efectos de algunos parámetros de las mediciones, como el intervalo de muestreo. Con estas características, el IRI fue adquiriendo una base de usuarios cada vez más amplia.

El cálculo del IRI se basa en un modelo matemático conocido como *modelo de cuarto de carro*, el cual representa un vehículo mediante un solo sistema rueda-suspensión, con la porción de la carrocería que le corresponde, como se muestra en la Figura 1.1 (Romero & Lozano, 1995). De acuerdo con la misma figura, el modelo está constituido por las masas del vehículo y la suspensión, un resorte vinculado a la masa del vehículo y uno más que representa a la rueda, así como un amortiguador.

De manera general, el IRI se calcula como el cociente entre el desplazamiento vertical (en valor absoluto) de la masa de la carrocería con respecto a la masa de la suspensión, dividido entre la distancia recorrida. De este modo, queda expresado en unidades como m/km, mm/m o in/mi. Esta acumulación del desplazamiento vertical es análoga a la que experimentan los pasajeros de un vehículo real y se relaciona directamente con la comodidad del viaje.

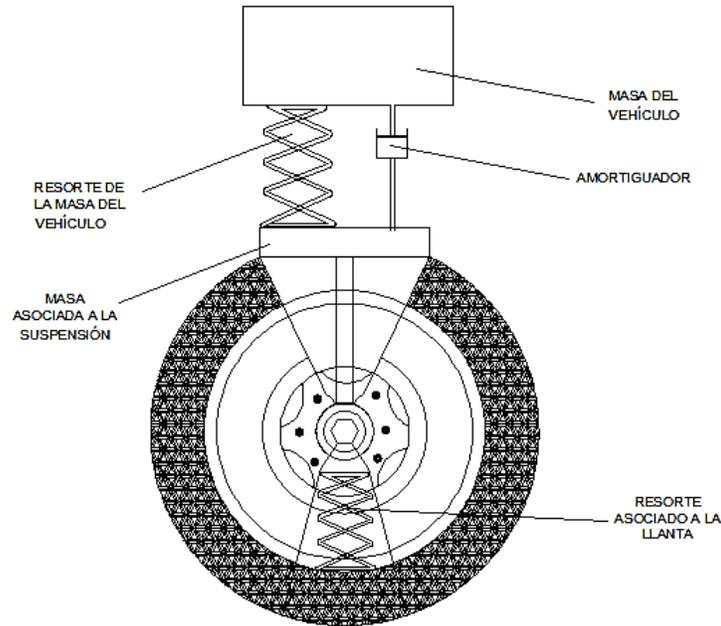


Figura 1.1 Modelo de cuarto de carro (Romero & Lozano, 1995).

El modelo de cuarto de carro fue ajustado para establecer una correlación entre los equipos de medición de tipo respuesta y el IRI calculado con el algoritmo propuesto por el Banco Mundial (Sayers, et al., 1986-1), el cual utiliza como dato de entrada el perfil longitudinal del tramo obtenido con un levantamiento topográfico.

Por lo que respecta a la escala del IRI, el valor mínimo es de cero, y corresponde a una superficie ideal perfectamente uniforme. En cuanto al valor máximo, no existe propiamente un límite, sin embargo, se considera que un camino con un IRI de 12 m/km o 760 in/mi es prácticamente intransitable. En la Figura 1.2 se muestra la escala para la medición del IRI originalmente propuesta por el Banco Mundial (Sayers, et al., 1986-1).

Cabe hacer notar que, en una superficie sin deformaciones con una pendiente constante, esto es, en un plano inclinado perfecto, el IRI es igual a cero. Así, la pendiente longitudinal de una carretera no influye en el valor del IRI, el cual depende más bien de los cambios locales de pendiente capaces de generar los desplazamientos entre la carrocería y la suspensión antes mencionados.

1.3 Procedimiento de análisis

En el año 2018, el IMT finalizó un estudio cuyo objetivo general consistió en determinar si la información histórica sobre el IRI recopilada por la SCT podría resultar útil para la obtención de modelos de deterioro a nivel de red y, de ser el caso, proponer una metodología para el desarrollo de estos modelos (Solorio, 2018).

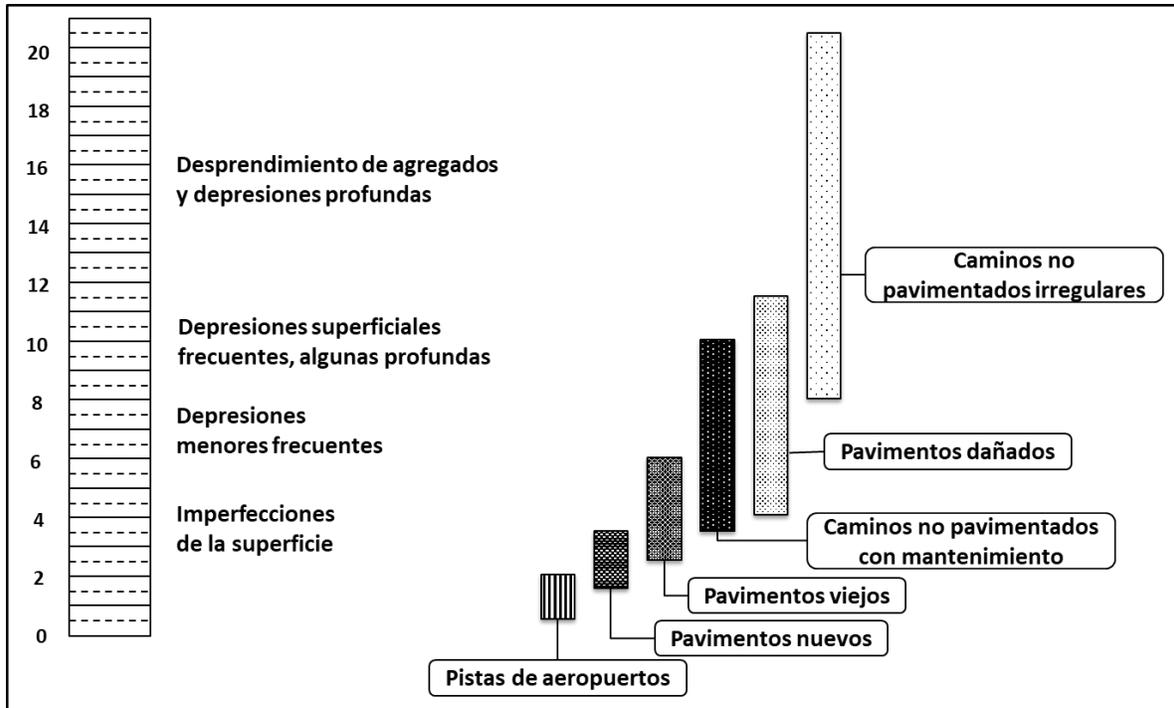


Figura 1.2 Escala de Valores del IRI y características de los pavimentos.

En el estudio se utilizó una muestra con datos de IRI de tres carreteras de cuota, que cubren una longitud total de 473.104 km-carril. Del análisis de esta muestra se concluyó que solo una parte de la información (alrededor del 37 %) era útil para la modelización del deterioro, ya que los datos examinados contenían una cantidad considerable de eventos de disminución del IRI entre dos años consecutivos que, evidentemente, representan una tendencia contraria al deterioro del pavimento. Los datos utilizados abarcan el periodo comprendido entre los años 2004 y 2013.

La metodología de modelización propuesta consta de los siguientes pasos:

- a) Definir los tramos de interés y clasificarlos en *familias*, con base en información complementaria como la estructura del pavimento, el tránsito, variables climáticas, etc. La metodología prevé el desarrollo de un modelo de deterioro por cada familia.
- b) Subdividir las familias en *grupos* en función del historial de reparaciones de los diferentes tramos.
- c) Para cada grupo:
 - i. Preparar series históricas de datos del IRI. En la literatura técnica aplicable, se recomienda que las series abarquen un periodo de por lo menos diez años (Butt, et al., 1994).
 - ii. Aplicar la *prueba de los signos* a las diferencias de IRI de cada par de mediciones anuales consecutivas, e identificar los pares de años en

los que se presentan diferencias negativas estadísticamente significativas. El segundo de cada uno de estos pares de años se considera el inicio de un ciclo de deterioro, esto es, un periodo delimitado por la ejecución de trabajos de conservación que mejoran el IRI¹.

- iii. El conjunto de datos útiles para la modelización estará constituido por los datos del ciclo de deterioro que contenga el año en el que fue recopilada la información del ciclo anterior. Si ninguno de los ciclos identificados cumple con esta condición, se considerará que no existen datos útiles para el grupo.
- d) Unir los conjuntos de datos útiles de los diferentes grupos de cada familia, a fin de obtener modelos de deterioro a este nivel.
- e) A partir de los datos de las familias, integrar los modelos de deterioro utilizando el procedimiento de cadenas Márkov.

En los siguientes incisos se describen brevemente los procesos de preparación de los datos, identificación de ciclos de deterioro y obtención de cadenas de Márkov.

1.4 Clasificación de los tramos

Los tramos de los que se tomó la información de IRI empleada en el estudio se clasificaron en función de los siguientes parámetros: espesor de carpeta o losa de concreto, deflexión central, tránsito diario promedio anual, precipitación media mensual y temperatura media anual. Para los espesores, las deflexiones y el tránsito, se consideraron valores medidos en 2008.

La clasificación se basó en tres niveles de los diferentes parámetros, identificados con los descriptores *Bajo(a)*, *Medio(a)* y *Alto(a)* (ver Tabla 1.1). Adicionalmente, se hizo una subdivisión de acuerdo con el tipo de pavimento de los tramos: asfáltico o de concreto hidráulico.

Tabla 1.1 Rangos de clasificación por parámetro de los tramos.

Parámetro	Rangos de valores por descriptor		
	<i>Bajo(a)</i>	<i>Medio(a)</i>	<i>Alto(a)</i>
Espesor de carpeta o losa (mm)	< 75	75 – 150	> 150
Deflexión central (mm)	<0.35	0.35 – 0.70	> 0.70
TDPA (veh/día)	5,000	5,000 – 15,000	> 15,000
Precipitación media mensual (mm)	50	50 – 125	> 125
Temperatura media anual (° C)	14	14 – 22	> 22

¹ Se ha observado que, con frecuencia, la reducción del IRI entre dos años consecutivos se debe a errores de medición. Sin embargo, dado que, de cualquier manera, estas mediciones no pueden emplearse en la modelización del deterioro, los datos se tratan como si reflejaran la ejecución de trabajos de conservación.

A partir de los descriptores de la tabla anterior, para identificar a las diferentes familias se definió un código de la forma:

[00|20]-*ECd-DCd-TDd-PMd-TMd*

donde:

[00|20]: Código del tipo de pavimento, 00 para pavimento asfáltico y 20 para pavimento de concreto hidráulico.

EC: Espesor de carpeta o losa.

DC: Deflexión central.

TD: Tránsito diario promedio anual.

PM: Precipitación media mensual.

TM: Temperatura media anual.

d: Inicial del descriptor que se asigna a cada parámetro dependiendo de su magnitud: *b* para bajo(a), *m* para medio(a) y *a* para alto(a).

Al concluir el proceso de clasificación, se encontró que todos los tramos de la muestra quedaron comprendidos en los rangos de tránsito bajo y temperatura media anual alta, de manera que estos parámetros ya no se tomaron en cuenta en el análisis y se excluyeron de los códigos antes definidos.

Para considerar la subdivisión en grupos de las familias según el historial de conservación, se agregó al código una terna formada por los años de ejecución de los últimos trabajos de reconstrucción, rehabilitación² y tratamiento superficial. Asimismo, se agregó un número consecutivo en los casos en los que se detectaron historiales diferentes previos a los últimos trabajos. En la Tabla 1.2 se muestra una relación de las familias y grupos resultantes del proceso de clasificación, los cuales constituyeron la fuente de datos para el proyecto original de desarrollo de los modelos de deterioro y para las pruebas con el software que se presenta en este documento.

1.5 Identificación de ciclos de deterioro

La viabilidad del procedimiento de análisis propuesto para el desarrollo de los modelos de evolución del Índice de Regularidad Internacional, depende de que la información histórica empleada como base contenga ciclos de deterioro útiles, lo cual queda determinado por dos factores: i) Que los datos describan uno o más procesos de degradación del pavimento caracterizados por el incremento continuo del IRI en periodos de dos o más años; ii) Que los datos complementarios utilizados

² En este contexto, se entiende como *rehabilitación* la realización de obras de refuerzo estructural del pavimento, ya sea mediante la colocación de una sobrecarpeta o mediante el fresado y reemplazo de la carpeta existente.

para la clasificación de los tramos correspondan a un año inscrito en alguno de los periodos antes mencionados.

Tabla 1.2 Resultado de la clasificación de los tramos en familias y grupos.

No.	Familia	Grupo	No.	Familia	Grupo
1	00-ECa-DCb-PMm	1995-2010-2010	9	00-ECm-DCb-PMa	2002-2007-2010
		1995-2012-2012			2002-2002-2006
2	00-ECa-DCm-PMa	2000-2012-2012			2002-2002-2010
		2003-2012-2012			2009-2009-2009 (1)
3	00-ECa-DCm-PMm	1995-2010-2010			2009-2009-2009 (2)
		2009-2009-2011	10	00-ECm-DCb-PMm	1995-2010-2010
		2009-2009-2009			1995-2012-2012
4	00-ECb-DCb-PMa	1993-2009-2012			2009-2009-2009
		1993-2010-2012 (2)			2002-2007-2007
5	00-ECb-DCb-PMm	1995-2012-2012			2002-2012-2012
		1993-2009-2012	11	00-ECm-DCm-PMa	2009-2009-2011
6	00-ECb-DCm-PMa	2009-2009-2011			1993-1993-1993
		1993-2010-2012			1993-1993-2003
		2005-2010-2012			1993-2007-2007
7	00-ECb-DCm-PMm	1995-2010-2010			2002-2012-2012
		1993-2009-2012			2002-2002-2002
8	00-ECm-DCa-PMa	2002-2007-2010			2002-2011-2011
		2002-2007-2007		00-ECm-DCm-PMa	2002-2002-2010
9	00-ECm-DCb-PMa	1995-2009-2011			2009-2009-2009 (1)
		2009-2009-2009			2009-2009-2009 (2)
		2002-2012-2012			2009-2009-2009 (3)
		2002-2011-2011 (1)			2009-2009-2009 (4)
		2002-2011-2011 (2)			2010-2010-2010
9	00-ECm-DCb-PMa	2002-2002-2002			2012-2012-2012
		2010-2010-2010 (1)	12	00-ECm-DCm-PMm	1995-1995-2005
		2010-2010-2010 (2)			2009-2009-2009
		2002-2007-2007	13	20-ECa-DCb-PMa	2000-2010

En rigor, los años que definen los ciclos de deterioro deben definirse, en primera instancia, a partir de la fecha de apertura inicial del tramo al tránsito y, enseguida, de las de cada uno de los trabajos subsecuentes de conservación periódica o reconstrucción.

Sin embargo, a menudo, los registros de obra que mantienen las organizaciones de carreteras contienen omisiones o imprecisiones que, aunadas al hecho de que estos registros usualmente se procesan en forma manual, complican la identificación de los ciclos. Adicionalmente, cualquier error de medición que produzca una disminución aparente del IRI, resulta equivalente a una acción de conservación, lo que obliga a la búsqueda de un método alternativo para la delimitación de cada ciclo. A este respecto, el procedimiento de modelización desarrollado en el IMT propone el uso de técnicas de inferencia estadística.

El posible sentido negativo de la variación entre dos series anuales consecutivas de mediciones de IRI, puede determinarse aplicando la *prueba t* para evaluar una

hipótesis sobre la diferencia entre las medias μ_1 y μ_2 de dos poblaciones. Para ello, la hipótesis se formularía como:

$$H_0: \mu_1 - \mu_2 = 0 \quad (1.1)$$

$$H_1: \mu_1 - \mu_2 < 0 \quad (1.2)$$

donde H_0 y H_1 representan la hipótesis nula y la hipótesis alternativa de la prueba, con la cual se desea identificar, específicamente, los casos en los que exista una diferencia negativa estadísticamente significativa entre las mediciones consecutivas.

Para que la aplicación de la prueba t sea válida, es necesario que se cumpla lo siguiente:

- Que las observaciones en las dos poblaciones sean independientes.
- Que los datos de las poblaciones tengan una distribución normal.

Evidentemente, la primera condición no se cumple para los datos de IRI, ya que, en ausencia de trabajos de conservación, las mediciones de un año dado están relacionadas con las del año anterior por haberse efectuado ambas en el mismo tramo. No obstante, en este caso es aplicable una variante conocida como *prueba t por pares*, la cual exige que los datos se recolecten en parejas cuyos elementos difieran por alguna característica. Los datos del IRI permiten cumplir con ambos requisitos, ya que, por una parte, pueden compararse por pares de mediciones anuales consecutivas, y por otra, estas varían en función del nivel de deterioro alcanzado por cada tramo en cada año.

En la prueba t por pares, la condición de normalidad no se aplica a las poblaciones individuales sino a las diferencias entre los pares. A fin de verificar esta condición, se aplicó la prueba *Anderson-Darling* a los datos mencionados en el inciso 1.3, con lo cual pudo verificarse que las diferencias anuales en las mediciones de IRI no se ajustan a la distribución normal. Este hallazgo condujo a la evaluación de métodos *no paramétricos* de inferencia estadística, cuya característica distintiva es que no se basan en distribuciones de probabilidad definidas por un cierto número de parámetros.

Entre los métodos no paramétricos, la *prueba de los signos* destaca porque no tiene ningún requisito relacionado con la distribución de los datos y, de hecho, se recomienda como sustituto de la prueba t por pares para muestras que no cumplen con la condición de normalidad (Buthmann, 2015).

A diferencia de la prueba t, las evaluaciones con la prueba de los signos no se realizan sobre las medias sino sobre las medianas de las poblaciones de interés. Así, para analizar estadísticamente el comportamiento del IRI en el tiempo, la prueba de hipótesis debe reformularse en los siguientes términos:

$$H_0: \eta_{dif} = 0 \quad (1.3)$$

$$H_1: \eta_{dif} < 0 \quad (1.4)$$

En las expresiones anteriores, η_{dif} se refiere a la diferencia entre las medianas de las mediciones consecutivas de IRI.

En general, dentro del trabajo inicial de desarrollo de los modelos de evolución del IRI, la aplicación de la prueba de los signos produjo resultados satisfactorios en lo que respecta a la identificación de pares de mediciones sugerentes de un descenso del IRI, lo que permitió la delimitación de ciclos de deterioro y la extracción de datos útiles para la modelización.

Sin embargo, en diversos casos, se pudo observar la aparente ocurrencia de errores tipo II, los cuales se refieren a la incapacidad de una prueba para rechazar la hipótesis nula cuando esta es falsa. Para ilustrar lo anterior, considérese el gráfico de la Figura 1.3 (Solorio, 2018).

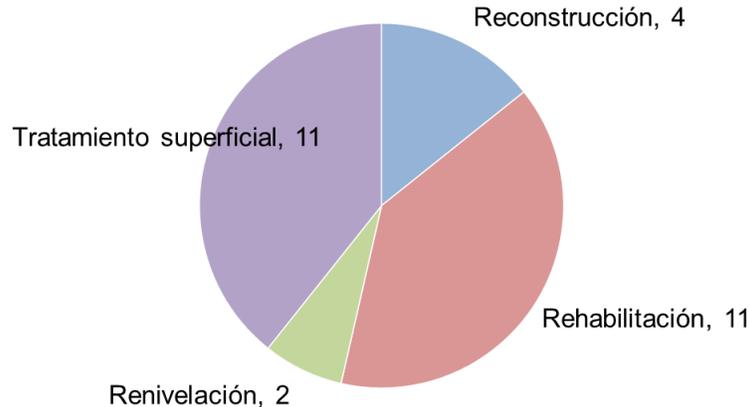


Figura 1.3 Prevalencia de la hipótesis nula en presencia de obra.

En este gráfico se muestra la clasificación por tipo de obra de casos en los que la prueba de los signos no rechazó la hipótesis nula aun cuando existían registros de ejecución de acciones de conservación. Como se aprecia en la figura, los casos referidos suman 28, lo que representa el 18 % del total de pruebas aplicadas (156). Si bien es cierto que la mitad de estos casos corresponden a tratamientos superficiales y renivelaciones, que normalmente no mejoran la regularidad, la otra mitad comprende trabajos de rehabilitación y reconstrucción que tienen, entre otros propósitos, lograr una disminución del IRI. Es en estos casos específicos que la prueba de los signos podría haber generado errores tipo II.

A la capacidad de una prueba para no incurrir en errores tipo II se le conoce como *potencia* y, precisamente, una de las desventajas de la prueba de los signos con respecto a la prueba t es su menor potencia. Por esta razón y, como parte de este trabajo, se decidió investigar con mayor detalle la aplicación de la prueba t por pares para la identificación de los ciclos de deterioro, como se explica en el capítulo 2.

1.6 Cadenas de Márkov

Existen dos métodos para el desarrollo de modelos de deterioro de pavimentos a partir de información histórica (también llamados modelos empíricos): el análisis de regresión y las cadenas de Márkov (Robinson, et al., 1998). Los modelos de regresión son deterministas, es decir, su aplicación produce un resultado único (Lytton, 1987). En cambio, las cadenas de Márkov tienen un carácter estocástico, de manera que simulan el deterioro empleando una distribución de probabilidad. Dado que este enfoque es más compatible con la incertidumbre inevitablemente vinculada a los datos sobre el estado del pavimento (Ningyuan, et al., 1996), en el proyecto de modelización del IMT se decidió utilizar el método de las cadenas de Márkov.

Para simularse con este método, el proceso de interés debe tener las siguientes características (Ortiz-García, et al., 2006): i) Ser discreto en el tiempo; ii) Tener un número finito de posibles estados; iii) Que el estado del proceso en el futuro dependa únicamente del estado actual y no de estados previos. Matemáticamente, una cadena de Márkov se define como:

$$\mathbf{a}_t = \mathbf{a}_0 \mathbf{P}^t \quad (1.5)$$

En la ecuación anterior, \mathbf{a}_t representa el *vector de estado* en el tiempo t , \mathbf{a}_0 el *vector de estado inicial* y \mathbf{P} la *matriz de probabilidades de transición (MPT)*.

Los vectores de estado pueden utilizarse para representar la condición del pavimento, si esta se expresa como una serie porcentajes de la longitud total que se encuentran en determinados estados predefinidos y asociados con rangos del indicador de deterioro de interés (p. ej. $1.8 \leq IRI < 2.5$ m/km).

Las MPT agrupan las probabilidades p_{ij} de que un segmento de carretera que se encuentra en el estado i se deteriore hasta el estado j durante un periodo de tiempo t (normalmente un año). Las cadenas de Márkov basadas en una sola MPT, como la definida de manera genérica por la ecuación 1.5, se denominan *cadenas homogéneas*.

En la Figura 1.4 (a) se muestra la forma general de una MPT. La Figura 1.4 (b) presenta una forma simplificada de esta matriz empleada a menudo en la modelización del deterioro, la cual toma en cuenta lo siguiente (Butt, et al., 1994): i) El deterioro no puede revertirse a menos que se ejecute una acción de conservación; ii) El pavimento solo puede cambiar de un estado de deterioro al siguiente durante un año. En el estudio del IMT se utilizó la forma simplificada de la matriz.

Existen varios procedimientos para calcular los elementos de las MPT. En este proyecto se utilizan los siguientes métodos:

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1n} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2n} \\ p_{31} & p_{32} & p_{33} & \cdots & p_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & p_{n3} & \cdots & p_{nn} \end{bmatrix} \quad (a)$$

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & 0 & \cdots & 0 \\ 0 & p_{22} & p_{23} & \cdots & 0 \\ 0 & 0 & p_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \quad (b)$$

Figura 1.4 Matrices de probabilidades de transición.

- a) Calcular las proporciones de los tramos que experimentan los cambios previstos en la MPT, utilizando la ecuación:

$$p_{ij} = \frac{N_{ij}}{N_i} \quad (1.6)$$

donde:

N_{ij} : Número de tramos que pasaron del estado i al estado j durante un año determinado.

N_i : Total de tramos que se encontraban en el estado i al inicio del año.

Estas proporciones se calculan entre pares de años consecutivos, por lo que, para obtener probabilidades de transición representativas de todo el periodo de análisis, deben promediarse las probabilidades calculadas para cada par, con lo que se obtiene una *MPT promedio*.

- b) Minimizar las diferencias entre los vectores de estado producidos por las cadenas de Márkov y por los datos disponibles, aplicando el proceso de optimización definido como:

$$Z = \min \sum_{t=1}^N \sum_{i=1}^M [a_t(i) - a'_t(i)]^2 \quad (1.7)$$

donde:

N : Años del periodo de análisis.

M : Número de estados de condición.

$a_t(i)$: Componentes de los vectores de estado obtenidos con las cadenas de Márkov para el año t .

$a'_t(i)$: Componentes de los vectores de estado obtenidos a partir de los datos originales para el mismo año.

- c) Minimizar las diferencias entre los valores esperados del IRI calculados con las cadenas de Márkov y los promedios de las mediciones efectuadas en el tramo. Con este objetivo, el problema de optimización puede plantearse como:

$$Z = \min \sum_{t=1}^N \{E[X(t)] - Y(t)\}^2 \quad (1.8)$$

En la expresión anterior, N tiene el mismo significado que en la ecuación 1.7. Adicionalmente:

$E[X(t)]$: Valor esperado del indicador de deterioro estimado con las cadenas de Márkov.

$Y(t)$: Promedio de las mediciones realizadas en el tramo en el año t .

En este trabajo, los procesos de optimización anteriores se llevaron a cabo empleando el *algoritmo de optimización no lineal del gradiente reducido generalizado*, disponible en el paquete de herramientas *Solver* de Excel (Fylstra, et al., 1998).

Los rangos a los que se asocian los estados predefinidos de la condición del pavimento, se pueden representar como:

$$\mathbf{R} = \{[x_1, x_2), [x_2, x_3), \dots, [x_{r-1}, x_r]\} \quad (1.9)$$

donde:

$$x_{k+1} > x_k \quad \forall k$$

Dado el conjunto de rangos \mathbf{R} , el valor esperado del indicador de deterioro para un ciclo de deterioro determinado puede calcularse mediante la expresión:

$$x_t = \mathbf{a}_t \cdot \mathbf{c} \quad (1.10)$$

en la que \mathbf{c} es el vector de marcas de clase de los rangos del indicador, es decir:

$$\mathbf{c} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_r) \quad (1.11)$$

2. Análisis estadístico

Como se menciona en el capítulo previo, el procedimiento de modelización del deterioro propuesto por el IMT se basa en la aplicación de técnicas de inferencia estadística para delimitar ciclos de degradación continua del pavimento o *ciclos de deterioro* que, en términos del IRI, se caracterizan por un incremento sostenido de esta medida de la regularidad.

En el estudio del Instituto al que se ha hecho referencia (Solorio, 2018), se optó por la prueba de los signos para determinar el final de cada uno de estos ciclos por la detección de un cambio negativo del IRI estadísticamente significativo entre dos series anuales consecutivas. Sin embargo, la sospecha respecto a la generación de errores tipo II con esta prueba, motivó el interés por hacer una investigación adicional de la prueba t por pares, cuya capacidad para evitar esa clase de errores es mayor. El objetivo principal de esta investigación consistió en determinar los posibles efectos negativos derivados del incumplimiento de los datos con el supuesto de normalidad de la prueba t.

En lo que sigue, se presentan descripciones más completas de las pruebas de los signos y t por pares. A la descripción de esta última, se añade el análisis de sus posibilidades de aplicación en la delimitación de los ciclos de deterioro. Por otro lado, a fin de poner en contexto los posibles errores que estarían generándose con la prueba de los signos, se examinan los dos tipos de errores que pueden presentarse en el contraste de hipótesis. Finalmente, se revisa el concepto de *valor P*, el cual se utiliza en este trabajo como resultado principal de las pruebas estadísticas.

2.1 Prueba de los signos

En la documentación del software *Dataplot* del NIST, publicada en línea, se establece que la prueba t puede considerarse como la “prueba estándar” para comparar dos poblaciones a través de la evaluación de hipótesis sobre sus medias (NIST, 2015). Sin embargo, se advierte que, si las poblaciones no tienen una distribución normal, la prueba t podría no ser válida, especialmente si las muestras son pequeñas.

El mismo sitio recomienda utilizar la prueba de los signos cuando se sospecha que los datos no cumplen con el supuesto de normalidad. No obstante, destaca el hecho de que esta prueba no tiene una potencia tan alta como la de la prueba t cuando los datos sí se distribuyen normalmente. Por otro lado, subraya que, por lo regular, la prueba de los signos compara las poblaciones en términos de las medianas y no de las medias.

Como ya se ha mencionado, cuando las poblaciones objeto del análisis son dependientes se utiliza la versión *por pares* de la prueba t. En este caso, las poblaciones corresponden al examen de los mismos objetos dos veces, ya sea porque se han hecho mediciones en ellos en diferentes momentos o porque han recibido dos tratamientos distintos. Claramente, lo anterior se corresponde con el análisis de mediciones del IRI realizadas en dos años consecutivos, es decir, con el problema que se aborda en este trabajo. Aunque la prueba de los signos se utiliza más comúnmente para el análisis de poblaciones por pares, también es posible emplearla en la evaluación de hipótesis para una sola muestra.

Si η_1 y η_2 representan las medianas de las muestras disponibles, la comparación de las poblaciones respectivas mediante la prueba de los signos puede efectuarse evaluando la hipótesis nula:

$$H_0: \eta_1 = \eta_2$$

en contraste con alguna de las siguientes hipótesis alternativas:

$$H_1: \eta_1 \neq \eta_2$$

$$H_1: \eta_1 < \eta_2$$

$$H_1: \eta_1 > \eta_2$$

La aplicación de la prueba de los signos se basa en el cálculo de las diferencias:

$$D_i = X_{1_i} - X_{2_i} \tag{2.1}$$

donde:

D_i : Diferencia entre dos elementos de las muestras bajo análisis.

X_{1_i} : Elemento de la primera muestra.

X_{2_i} : Elemento de la segunda muestra.

Cabe enfatizar que las parejas (X_{1_i}, X_{2_i}) que se utilizan en el cálculo de las diferencias deben estar referidas a los mismos objetos, o segmentos de carretera en el caso particular que se trata en este documento.

Considérese que $R+$ y $R-$ representan, en forma respectiva, el número de veces que la diferencia D_i entre dos observaciones resulta positiva o negativa. Si las muestras tuvieran medianas iguales y fueran simétricas, entonces $R+$ y $R-$ deberían ser similares, y podría concluirse que no existe evidencia suficiente para rechazar la hipótesis nula. Por el contrario, si se identificaran valores excesivos de $R+$ o $R-$, entonces esta hipótesis tendría que rechazarse. Nótese que las diferencias nulas no se toman en cuenta. Dado que solo se considera el carácter positivo o negativo de las diferencias, el estadístico de la prueba de los signos se distribuye en forma binomial con $p = 0.5$ (NIST, 2015).

En términos generales, los elementos del proceso de contraste de hipótesis basado en la prueba de los signos pueden definirse según lo indicado en la Tabla 2.1. En esta tabla n se refiere al tamaño de las muestras y p a la probabilidad del ensayo de Bernoulli de la distribución binomial ($p = 0.5$).

Tabla 2.1 Elementos de la prueba de los signos.

Elemento	Definición
Hipótesis nula (H_0)	$\eta_1 = \eta_2$
Hipótesis alternativas (H_1)	$\eta_1 \neq \eta_2$ $\eta_1 < \eta_2$ $\eta_1 > \eta_2$
Estadísticos	$S_{+=} P(X \leq R_{+}) = \sum_{k=0}^{R_{+}} \binom{n}{k} p^k (1-p)^{n-k}$ $S_{-} = P(X \leq R_{-}) = \sum_{k=0}^{R_{-}} \binom{n}{k} p^k (1-p)^{n-k}$
Nivel de significación α	Usualmente 0.05
Región crítica	$S_{+} < \alpha$, prueba unilateral: $\eta_1 < \eta_2$ $S_{-} < \alpha$, prueba unilateral: $\eta_1 > \eta_2$ $\frac{\alpha}{2} < S_{+} < 1 - \frac{\alpha}{2}$, prueba bilateral: $\eta_1 = \eta_2$
Conclusión	Rechazar la hipótesis nula si el estadístico de la cae en la región crítica

Adicionalmente, debe señalarse que, debido a que la distribución binomial es una función discreta, el nivel de significación real no será exacto en la mayoría de los casos. También, debe recordarse que la hipótesis alternativa útil para la delimitación de los ciclos de deterioro es $H_1: \eta_1 > \eta_2$, o $H_1: (\eta_2 - \eta_1) < 0$, es decir, la que implica una disminución del IRI entre las mediciones realizadas en dos años consecutivos.

2.2 Prueba t por pares

De manera simple y llana, la enciclopedia de la Sociedad Europea de Matemáticas define de a la *prueba de Student* o *prueba t* como una prueba de significación para el valor medio de una distribución normal (The European Mathematical Society, 2020).

De acuerdo con este mismo recurso, una *prueba de significación* es uno de los métodos básicos para la evaluación de hipótesis estadísticas, las cuales “se utilizan para confrontar los resultados de las observaciones $x_1 \dots x_n$, interpretadas como realizaciones de las variables aleatorias $X_1 \dots X_n$, con una hipótesis H_0 ”. Normalmente, este tipo de pruebas se basan en el valor observado de un estadístico apropiado $T = T(X_1 \dots X_n)$, cuya forma específica depende del planteamiento de cada problema.

En el sitio web del software Dataplot (NIST, 2018) se distinguen tres variantes de la prueba t, con las que se evalúan las siguientes hipótesis:

a) La media de una muestra es igual a un valor especificado:

$$H_0: \mu_1 = \langle \text{valor} \rangle$$

$$H_1: \mu_1 \neq \langle \text{valor} \rangle$$

b) Las medias de dos muestras independientes son iguales:

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

c) Las medias de dos muestras dependientes, o tomadas por pares, son iguales. Para esta variante, el planteamiento de las hipótesis es igual al del caso anterior.

La Tabla 2.2 resume el proceso de contraste de la última variante de la prueba t, a la cual suele denominarse *prueba t por pares*.

Tabla 2.2 Elementos de la prueba t por pares.

Elemento	Definición
Hipótesis nula (H_0)	$\mu_1 = \mu_2$
Hipótesis alternativas (H_1)	$\mu_1 \neq \mu_2$ $\mu_1 < \mu_2$ $\mu_1 > \mu_2$
Estadístico	$t_0 = \frac{\bar{D}}{S_D/\sqrt{n}}$, donde: $\bar{D} = \frac{\sum_{i=1}^n D_i}{n}$, $S_D^2 = \frac{\sum_{i=1}^n D_i^2 - [(\sum_{i=1}^n D_i)^2/n]}{n-1}$ y $D_i = X_{1i} - X_{2i}$
Nivel de significación α	Usualmente 0.05
Región crítica	$t_0 < t_{\alpha, n-1}$, prueba unilateral: $\mu_1 < \mu_2$ $t_0 > t_{1-\alpha, n-1}$, prueba unilateral: $\mu_1 > \mu_2$ $t_0 < t_{\alpha/2, n-1}$ o $t_0 > t_{1-\alpha/2, n-1}$, prueba bilateral: $\mu_1 \neq \mu_2$
Conclusión	Rechazar la hipótesis nula si el estadístico de la prueba cae en la región crítica

En la sección de la tabla referente al cálculo del estadístico, D_i es la diferencia entre los elementos de las muestras tomadas por pares y \bar{D} y S_D^2 la media y la varianza de estas diferencias, mientras que n también representa el tamaño de las muestras. Por otro lado, se enfatiza que la hipótesis alternativa útil para el cierre de un ciclo de deterioro es $H_1: \mu_1 > \mu_2$, o $H_1: (\mu_2 - \mu_1) < 0$.

Según lo esbozado en el inciso 1.5, durante la fase de desarrollo de la metodología de modelización se descartó el uso de la prueba t por pares y se optó por la prueba

de los signos debido a que las diferencias anuales en las mediciones de IRI no se ajustan a la distribución normal, lo cual constituye una condición para la aplicación de la prueba t. Sin embargo, dada la sospecha sobre la ocurrencia de errores tipo II con la prueba de los signos, lo cual sería una manifestación de la menor potencia atribuida a esta prueba, en esta nueva fase del trabajo se propuso reevaluar la aplicabilidad de la prueba t por pares.

Como parte de esta tarea, luego de una búsqueda exhaustiva, se encontró en Internet un documento del software estadístico Minitab aparentemente utilizado por la empresa desarrolladora para la impartición de un curso de capacitación. En este documento, cuyo título puede traducirse al español como *Inferencia estadística y pruebas t* (Minitab, 2010), se indica que, antes de usar la prueba t por pares, se debe verificar que:

- Las observaciones se hicieron efectivamente en pares.
- Los datos son continuos.
- Las diferencias se ajustan a una distribución normal.

Sin embargo, enseguida se afirma que la prueba t seguirá siendo una opción “bastante robusta” frente a violaciones del supuesto de normalidad siempre que las observaciones se hagan de manera aleatoria y que la distribución de los datos sea continua, unimodal y “razonablemente simétrica”.

Por otro lado, el documento destaca que la variante por pares de la prueba tiene una potencia mayor que la variante de dos muestras independientes, debido a que la dependencia entre las observaciones en el primer caso redundaba en una menor variabilidad del análisis.

Adicionalmente, en la sección “*Data Considerations for Paired t*” (“Consideraciones sobre los datos de la t por pares”) del sitio de asistencia técnica *Minitab Express Support* de Minitab (Minitab, 2019-1), se enumeran las siguientes directrices para asegurar la validez de los resultados obtenidos con esta prueba:

- Los datos deben ser continuos.
- Los datos no deben estar excesivamente sesgados y el tamaño de la muestra debe ser mayor que 20.
- Las observaciones deben haberse recopilado por pares.
- La muestra debe haberse elegido aleatoriamente.

Como puede apreciarse, estas directrices son compatibles con los enunciados del primer documento.

Al revisar la información sobre el IRI empleada en el presente trabajo a la luz de las consideraciones de los documentos de Minitab, se encontró que los datos disponibles cumplen con los requisitos relativos a los siguientes atributos: distribución continua y unimodal, tamaño de la muestra ($50 \leq n \leq 1917$ para los diferentes conjuntos), recopilación en pares y muestreo aleatorio.

Sin embargo, en lo que respecta a la simetría de las distribuciones se observó que, aunque en términos generales las distribuciones de probabilidad de las diferentes familias son “razonablemente simétricas”, existen algunos conjuntos en los que el sesgo parece ser excesivo. Es la condición que se muestra, a manera de ejemplo, en la Figura 2.1, en la que puede apreciarse la distribución de las diferencias entre los IRI medidos en 2005 y 2006 en los tramos del grupo 1993-2009-2012 de la familia 00-ECb-DCb-PMm.

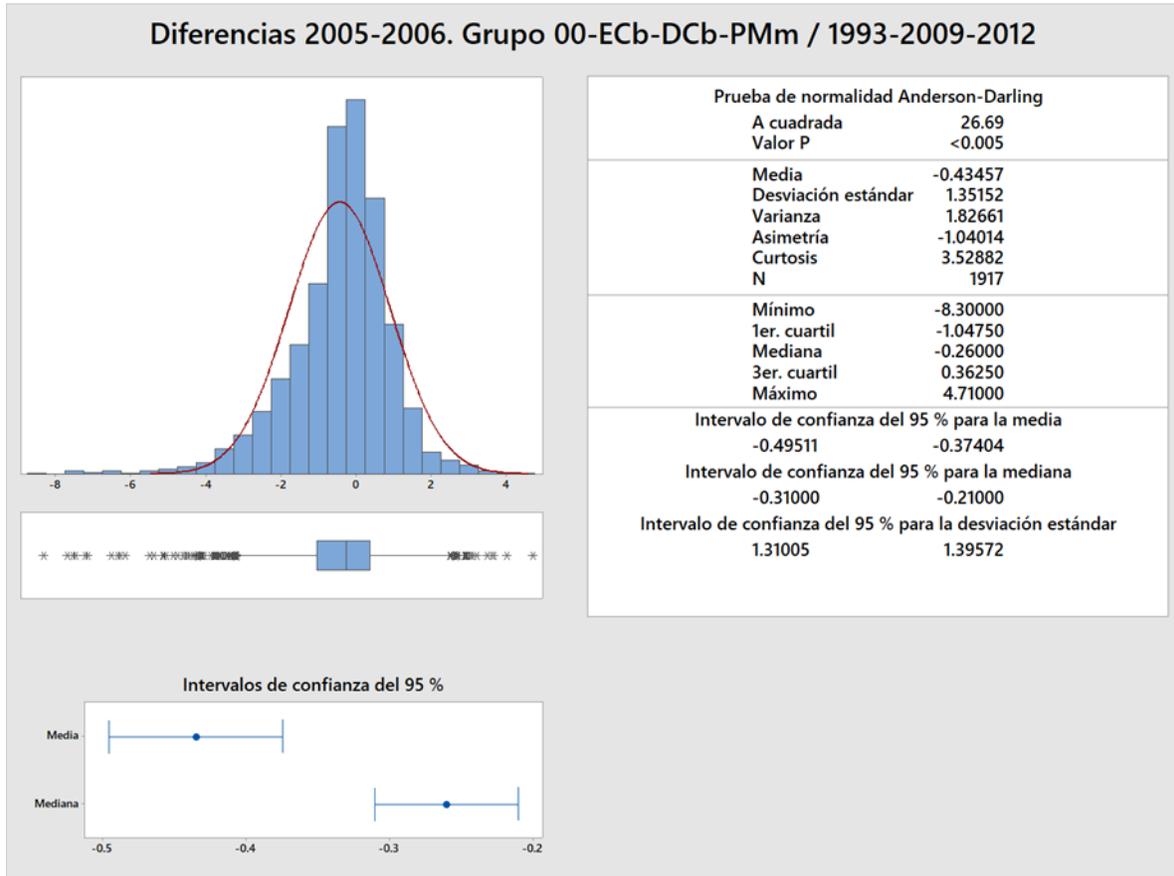


Figura 2.1 Ejemplo de la asimetría de las distribuciones de diferencias de IRI.

A fin de interpretar la propiedad de asimetría de manera objetiva, Brown (Brown, 2020-1) recomienda aplicar la siguiente regla general propuesta originalmente en (Bulmer, 1979):

- Si la asimetría es menor que -1 o mayor que +1, la distribución es altamente asimétrica.
- Si la asimetría se encuentra entre -1 y $-\frac{1}{2}$ o entre $+\frac{1}{2}$ y +1, la distribución es moderadamente asimétrica.
- Si la asimetría se encuentra entre $-\frac{1}{2}$ y $+\frac{1}{2}$, la distribución es aproximadamente simétrica.

Aplicando esta regla a la totalidad de las mediciones que corresponden al grupo *00-ECb-DCb-PMm / 1993-2009-2012*, se determinaron los niveles de asimetría por pares de mediciones consecutivas de la Tabla 2.3. De acuerdo con esta tabla, dos terceras partes de las series por pares del grupo mencionado tienen un nivel de asimetría de *bajo a moderado* y, de las tres series con un nivel alto, dos de ellas tienen un valor muy cercano al límite inferior.

Tabla 2.3 Niveles de asimetría en el grupo *ECb-DCb-PMm / 1993-2009-2012*.

Series anuales consecutivas	Asimetría	Nivel
2004-2005	-0.53	Moderada
2005-2006	-1.04	Alta
2006-2007	0.84	Moderada
2007-2008	0.67	Moderada
2008-2009	-1.45	Alta
2009-2010	1.05	Alta
2010-2011	0.38	Baja
2011-2012	-0.16	Baja
2012-2013	0.58	Moderada

El patrón descrito se repite de manera más o menos similar para el resto de los grupos y familias del paquete de datos, por lo que se vuelve imperativo explorar el uso de la prueba t por pares como método alternativo para determinar los límites de los ciclos de deterioro. Esta necesidad cobra especial relevancia ante la duda sobre la generación de errores tipo II con la prueba de los signos, cuya aplicación podría estar conduciendo, bajo ciertas condiciones, al uso de datos inapropiados para la modelización del deterioro.

Una vez verificada la posibilidad de aplicación de la prueba t en el análisis, se decidió añadir su codificación al desarrollo del software para la generación de modelos de evolución del IRI.

2.3 Errores en las pruebas de hipótesis

La aplicación de una prueba de hipótesis puede producir cualquiera de los resultados que se muestran en la Tabla 2.4 (Brown, 2020-2).

Tabla 2.4 Posibles resultados de una prueba de hipótesis.

	Rechazar H_0 , aceptar H_1	No rechazar H_0
H_0 es verdadera	Error tipo I	Resultado correcto
H_0 es falsa (y H_1 verdadera)	Resultado correcto	Error tipo II

Como se observa, entre los posibles resultados figuran dos tipos de errores. El primero de ellos, conocido como *error tipo I*, ocurre si se rechaza la hipótesis nula cuando es en realidad verdadera. Los errores tipo I se encuentran estrechamente vinculados con el nivel de significación de la prueba que, de hecho, puede definirse como la probabilidad de ocurrencia de esta clase de error, es decir (Montgomery & Runger, 2003):

$$\alpha = P(\text{error tipo I}) = P(\text{rechazar } H_0 \text{ cuando es verdadera})$$

En forma equivalente, el error tipo II puede también expresarse en términos de probabilidad como:

$$\beta = P(\text{error tipo II}) = P(\text{no rechazar } H_0 \text{ cuando es falsa})$$

En relación con los errores antes enunciados, Montgomery y Runger enfatizan lo siguiente (Montgomery & Runger, 2003):

- a) Dado que la probabilidad de ocurrencia de un error tipo I es constante e igual al nivel de significación α elegido para el análisis, el rechazo de la hipótesis nula H_0 se considera normalmente una conclusión sólida.
- b) Por el contrario, la probabilidad β de un error tipo II no es constante, sino que varía inversamente con respecto al nivel de α seleccionado, al tamaño de la muestra y a la cercanía entre el valor real del parámetro y el consignado en H_0 . Por esta razón, es usual juzgar la aceptación de la hipótesis nula como una conclusión débil, lo que ha llevado a que este resultado se enuncie empleando la frase “no rechazar H_0 ”, que implica que no se ha encontrado suficiente evidencia para rechazar esta hipótesis.

A partir de la probabilidad de generar un error tipo II, la llamada *potencia* de una prueba estadística se calcula como:

$$\text{Potencia} = 1 - \beta,$$

por lo que este concepto adicional puede definirse como la probabilidad de rechazar la hipótesis nula mediante una prueba estadística cuando la hipótesis alternativa es verdadera. En otras palabras, la potencia es una medida de la sensibilidad de la prueba, entendida como su capacidad para detectar diferencias.

En el contexto de la comparación de mediciones anuales consecutivas de IRI, la ocurrencia de los errores antes descritos pueden interpretarse como se indica a continuación:

Error tipo I. Descartar datos para la modelización del deterioro aun cuando existe una probabilidad igual al nivel de confianza (o $1 - \alpha$) de que los datos sean útiles.

Error tipo II. No descartar datos para la modelización del deterioro debido a que no existe suficiente evidencia estadística para ello.

Como se ha mencionado, la experiencia en la aplicación de la prueba de los signos acumulada durante el desarrollo de la metodología de modelización (Solorio, 2018) muestra que es posible que solo se hayan generado errores tipo II durante ese ejercicio.

2.4 Valor P

De acuerdo con las tablas 2.1 y 2.2, las regiones críticas de una prueba de hipótesis dependen de los estadísticos definidos para cada prueba y del nivel de significación α empleado en el análisis. La hipótesis nula se rechaza siempre que el estadístico queda ubicado en la región crítica, razón por la también que se conoce a esta región como *región de rechazo*. A los valores del estadístico que delimitan la región de rechazo se les denomina *valores críticos*.

Otra medida cuantitativa para expresar el resultado de una prueba estadística es el llamado *valor P*, el cual puede definirse como la probabilidad de que el valor del estadístico de la prueba sea al menos tan extremo como el que se observa cuando la hipótesis nula es verdadera (NIST/SEMATECH, 2013). Un valor P pequeño es siempre indicativo de que la hipótesis nula es falsa.

En general, se recomienda definir con antelación al análisis el valor P que se utilizará como referencia para rechazar la prueba, del mismo modo que se elige el nivel de significación α . Coincidentemente, un valor P de uso común es $P = 0.05$, esto es un valor idéntico al que suele emplearse para α . En cualquier caso, es muy importante tener presente la diferencia entre ambos conceptos, ya que varios programas de estadística (p. ej. Minitab) generan sus resultados en términos del valor P y no de los valores críticos. En lo que se refiere al presente estudio, desde el desarrollo de la metodología de modelización se utiliza también el valor P para expresar el resultado de pruebas de hipótesis.

El método de cálculo del valor P depende de cada prueba de hipótesis y de la hipótesis alternativa que se utilice. En el proyecto inicial se utilizó Minitab para la aplicación de todas las funciones de tipo estadístico; por esta razón, a continuación se presentan los métodos empleados por ese programa para la estimación del valor P.

En lo que respecta a la prueba de los signos, Minitab utiliza dos procedimientos. El primero es aplicable para $n \leq 50$ y el software hace referencia a él como “método exacto”. Si se cumple la condición especificada, Minitab utiliza la distribución binomial con una muestra de tamaño n , después de omitir cualesquier observaciones iguales al valor de la mediana considerado en la hipótesis nula, y una probabilidad de ocurrencia $p = 0.5$. La expresión que corresponde a la hipótesis alternativa $H_1: (\eta_2 - \eta_1) < 0$ es (Minitab, 2019-2):

$$P(S \leq s) \tag{2.2}$$

donde s es el número de datos mayores que la mediana de la hipótesis nula y S es una variable aleatoria que sigue una distribución binomial de n intentos y 0.5 de probabilidad de ocurrencia.

Cuando $n > 50$, Minitab utiliza la siguiente aproximación normal a la distribución binomial:

$$Z = \frac{s - \frac{n}{2}}{\sqrt{\frac{n}{4}}}$$

en la que Z se distribuye de manera aproximadamente normal con $\mu = 0$ y $\sigma = 1$. En la misma expresión S es el número de observaciones por encima de la mediana, y tiene una distribución binomial con número de intentos igual a n y probabilidad de éxito $p = 0.5$ bajo la hipótesis nula. La aproximación normal al valor P que corresponde a la hipótesis alternativa $H_1: (\eta_2 - \eta_1) < 0$ es:

$$P\left(Z \leq \frac{s + 0.5 - \frac{n}{2}}{\sqrt{\frac{n}{4}}}\right) \quad (2.3)$$

donde s tiene el mismo significado que en la ecuación 2.8. En la ecuación 2.9 se utiliza una corrección de continuidad de 0.5.

Por lo que toca a la prueba t por pares, en el caso de la hipótesis alternativa $H_1: (\mu_2 - \mu_1) < 0$, el valor P se calcula como (Minitab, 2019-3):

$$P(t_{n-1} \leq t_0 | \mu_D = 0) \quad (2.4)$$

donde:

μ_D : Media de la muestra de diferencias.

t_0 : Estadístico t de los datos.

t_{n-1} : Variable aleatoria de la distribución t con $n - 1$ grados de libertad.

n : Tamaño de la muestra.

3. Desarrollo del software

Se presenta en este capítulo una descripción general del desarrollo del software para la generación de modelos de evolución del IRI basados en información histórica y en cadenas de Márkov. Se abordan específicamente los siguientes aspectos: necesidades que dieron pie al desarrollo de la herramienta, aspectos de su diseño, plataforma informática para la generación del código y la ejecución del programa e implementación de los principales procesos, es decir, identificación de datos útiles y confección de los modelos propiamente dichos. Por último, se revisan algunas características de la interfaz de usuario.

3.1 Definición de requerimientos

El proyecto del que se derivan el software y este informe tiene su origen en la necesidad de automatizar las siguientes tareas del proceso de modelización de la evolución del IRI en el tiempo desarrollado en el IMT:

- a) Leer las series de mediciones anuales de IRI preparadas para el conjunto de tramos clasificados en familias y grupos según lo descrito en los incisos 1.3 y 1.4. Como se indica en el inciso 1.3, estas series abarcan el periodo 2004 - 2013. Los datos de las familias se encuentran en archivos separados de Excel (uno por cada familia y cada carretera), y tienen un formato predefinido que incluye la siguiente información, organizada en columnas:
 - i. Clave de cinco caracteres numéricos de la carretera a la que pertenece el tramo, definida por la Dirección General de Servicios Técnicos de la SCT.
 - ii. Cadenamientos inicial y final de las mediciones de IRI efectuadas a cada 20 m.
 - iii. Mediciones anuales de 2004 a 2013.
 - iv. Columna en blanco.
 - v. Diferencias de IRI entre mediciones de años consecutivos.
- b) Para cada juego de diferencias entre pares de mediciones anuales consecutivas de IRI de cada grupo y cada familia:
 - i. Calcular los siguientes parámetros: media, desviación estándar, mediana, incremento porcentual de la media e incremento porcentual de la desviación estándar.

- ii. Aplicar la prueba de los signos para determinar la significación estadística de diferencias negativas, utilizando las hipótesis definidas por las ecuaciones 1.3 y 1.4. Emplear como criterio de decisión el valor P de la prueba.
 - iii. Organizar los resultados de acuerdo con el arreglo ilustrado en la Tabla 3.1 con la familia *00-ECa-DCm-PMa* (ver inciso 1.4). Este ejemplo fue extraído de una serie de archivos de Excel creados durante el desarrollo de la metodología, y las operaciones de cálculo de los parámetros de estadística descriptiva y de aplicación de la prueba de los signos fueron realizados con el software Minitab.
 - iv. Identificar los conjuntos de datos útiles para la modelización o, en otras palabras, las series de mediciones anuales consecutivas sin diferencias negativas estadísticamente significativas. Estas series, debe incluir la medición del IRI para el año de registro de la información utilizada para clasificar los tramos que, en el caso el paquete de datos del proyecto, corresponde a 2008. El conjunto de datos útiles constituye un ciclo de deterioro como los descritos en el inciso 1.5.

La Tabla 3.1 muestra los parámetros estadísticos de los años con datos útiles resaltados en color verde. En esta misma tabla, puede apreciarse que el primer grupo de la familia tiene un periodo bienal de datos útiles más que la segunda, lo cual se traduce en una *discontinuidad* en los datos inducida por el cambio en el tamaño y características de la muestra que sirve de base para la modelización. Nótese que esta clase de discontinuidades pueden presentarse al inicio o al final del periodo cubierto por los datos útiles de cada grupo.
 - v. Realizar las acciones necesarias para replicar los pasos i a iv anteriores con la prueba t por pares.
- c) Haciendo uso de los conjuntos de datos útiles a que se refiere el inciso anterior (esto es, los generados con base en las dos pruebas estadísticas), obtener las matrices de probabilidades de transición (MPT) de cada familia con los métodos descritos por las ecuaciones 1.6, 1.7 y 1.8. Mediante un análisis de la distribución de los datos de IRI realizado durante el desarrollo de la metodología de modelización, se determinó la conveniencia de utilizar los rangos de clasificación del IRI de la Tabla 3.2.

En la Tabla 3.3 se presentan las MPT obtenidas durante la primera fase del proyecto para la misma familia *00-ECa-DCm-PMa*. Este, y los demás resultados relativos a los modelos basados en Cadenas de Márkov que se citan en los siguientes párrafos, se encuentran en un archivo adicional de Excel creado para almacenar el modelo que corresponde a esta familia.

Tabla 3.1 Aplicación de la prueba de los signos, familia 00-ECa-DCm-PMa.

	Grupo 2000 - 2012 -2012																	
	2004-2005	2005-2006	2006-2007	2007-2008	2008-2009	2009-2010	2010-2011	2011-2012	2012-2013	2004-2005	2005-2006	2006-2007	2007-2008	2008-2009	2009-2010	2010-2011	2011-2012	2012-2013
Media dif.	-0.071	-0.756	0.089	-0.054	0.111	0.357	-0.381	-0.270	0.947	0.304	-0.439	0.214	0.014	0.309	0.062	-0.202	0.175	-0.026
Desv. est. dif.	1.694	1.636	1.856	1.102	1.492	1.203	1.475	1.597	1.769	1.088	1.009	0.759	0.746	0.788	0.868	1.097	1.170	1.131
Mediana dif.	-0.095	-0.843	0.118	-0.063	0.200	0.063	-0.248	-0.085	0.775	0.250	-0.345	0.175	0.010	0.325	-0.145	-0.210	0.285	-0.040
Valor p	0.285	0.000	0.927	0.064	0.990	0.898	0.000	0.195	1.000	1.000	1.000	1.000	0.662	1.000	0.000	0.000	1.000	0.089
Inc. media	-2.2%	-23.7%	3.7%	-2.2%	4.5%	13.8%	-11.1%	-10.9%	40.7%	17.1%	-21.2%	13.1%	0.8%	16.6%	2.3%	-8.6%	8.6%	-1.2%
Inc. desv. est.	8.1%	-2.3%	12.3%	-14.8%	-29.6%	90.7%	-12.3%	-34.2%	71.7%	45.6%	-42.5%	14.0%	6.6%	-19.1%	89.1%	-1.6%	-18.2%	4.4%

Tabla 3.2 Rangos de clasificación del IRI para la generación de las MPT.

Estado de condición	1	2	3	4
Rango de IRI (m/km)	< 1.5	1.5 - 2.5	2.5 - 3.5	> 3.5

Tabla 3.3 Matrices de probabilidades de transición, familia 00-ECa-DCm-PMa.

		MPT promedio			
		< 1.50	2.50	3.50	> 3.50
1.50	0.20	0.80	0.00	0.00	
2.50	0.00	0.80	0.20	0.00	
3.50	0.00	0.00	0.68	0.32	
> 3.50	0.00	0.00	0.00	1.00	

		MPT opt. (vectores de estado)			
		< 1.50	2.50	3.50	> 3.50
1.50	0.76	0.24	0.00	0.00	
2.50	0.00	0.92	0.08	0.00	
3.50	0.00	0.00	0.72	0.28	
> 3.50	0.00	0.00	0.00	1.00	

		MPT opt. (valores esperados)			
		< 1.50	2.50	3.50	> 3.50
1.50	1.00	0.00	0.00	0.00	
2.50	0.00	1.00	0.00	0.00	
3.50	0.00	0.00	0.09	0.91	
> 3.50	0.00	0.00	0.00	1.00	

- d) Con las MPT obtenidas, generar las cadenas de Márkov empleando la ecuación 1.5. Calcular las diferencias de las componentes de los vectores de estado calculados y originales, así como la suma de errores al cuadrado que corresponden a cada año y el error total del periodo. Dar a los resultados un acomodo como el que se utiliza en la Tabla 3.4 para presentar las cadenas de la familia 00-ECa-DCm-PMa.

En las cadenas de la Tabla 3.4 aparecen dos filas para el año 2009 que, además, han sido resaltadas en color amarillo. Esto último no se debe a la discontinuidad señalada en párrafos previos, sino al hecho de que, en general, las mediciones efectuadas ese año parecen haber sido objeto de un filtrado anormal que eliminó los valores extremos, de manera que estas mediciones resultan atípicas al compararlas con las del resto del periodo, como puede observarse en la Figura 3.1. Por esta razón, en el proceso original de modelización se decidió abrir la posibilidad de tomar en cuenta o no los datos de 2009. Los encabezados $\sum err_1^2$ y $\sum err_2^2$ que aparecen en las tablas, corresponden a estas dos opciones.

- e) Con base en las ecuaciones 1.10 y 1.11, calcular los valores esperados de IRI a partir de las tres versiones de las cadenas de Márkov, así como las diferencias con los valores correspondientes a las mediciones y las sumas de errores al cuadrado, con y sin los datos de 2009 si es que están incluidos en el ciclo de deterioro. Organizar el cálculo siguiendo el formato de la Tabla 3.5.
- f) Representar la evolución en el tiempo de los valores esperados de IRI en un gráfico como el de la Figura 3.2.

Tabla 3.4 Cadenas de Márkov, familia 00-ECa-DCm-PMa.

MPT promedio														
Año	Vectores de estado calculados			Vectores de estado originales			Diferencias			Σerr^2_1	Σerr^2_2			
	< 1.50	2.50	3.50	< 1.50	2.50	3.50	< 1.50	2.50	3.50			> 3.50		
2006	0.38	0.50	0.09	0.03	0.38	0.50	0.09	0.03				0.05	0.05	
2007	0.08	0.70	0.16	0.06	0.26	0.57	0.13	0.04	-0.18	0.12	0.03	0.02	0.08	
2008	0.02	0.62	0.25	0.11	0.25	0.58	0.13	0.04	-0.23	0.04	0.12	0.07	0.08	
2009	0.00	0.51	0.30	0.19	0.03	0.72	0.21	0.05	-0.02	-0.21	0.09	0.14	0.07	0.00
2009	0.04	0.50	0.32	0.15	0.04	0.50	0.32	0.15	0.00	0.00	0.00	0.00	0.00	0.00
2010	0.01	0.42	0.32	0.25	0.06	0.45	0.24	0.25	-0.05	-0.03	0.08	0.00	0.01	0.01
2011	0.00	0.34	0.30	0.35									0.21	0.14
2012	0.00	0.27	0.28	0.45										
2013	0.00	0.22	0.25	0.53										

MPT optimizada (vectores de estado)														
Año	Vectores de estado calculados			Vectores de estado originales			Diferencias			Σerr^2_1	Σerr^2_2			
	< 1.50	2.50	3.50	< 1.50	2.50	3.50	< 1.50	2.50	3.50			> 3.50		
2006	0.38	0.50	0.09	0.03	0.38	0.50	0.09	0.03					0.00	0.00
2007	0.29	0.55	0.10	0.06	0.26	0.57	0.13	0.04	0.03	-0.02	-0.02	0.02	0.00	0.00
2008	0.22	0.58	0.12	0.09	0.25	0.58	0.13	0.04	-0.03	0.00	-0.02	0.05	0.00	0.00
2009	0.16	0.59	0.13	0.12	0.03	0.72	0.21	0.05	0.14	-0.13	-0.08	0.07	0.05	0.00
2009	0.04	0.50	0.32	0.15	0.04	0.50	0.32	0.15	0.00	0.00	0.00	0.00	0.00	0.00
2010	0.03	0.47	0.27	0.24	0.06	0.45	0.24	0.25	-0.03	0.02	0.03	-0.01	0.00	0.00
2011	0.02	0.44	0.23	0.31									0.05	0.01
2012	0.02	0.41	0.20	0.37										
2013	0.01	0.38	0.18	0.43										

Tabla 3.4 Cadenas de Márkov, familia 00-ECa-DCm-PMa (cont.).

Año	MPT optimizada (valores esperados)													Σerr^2_1	Σerr^2_2		
	Vectores de estado calculados						Vectores de estado originales						Diferencias				
	< 1.50	2.50	3.50	> 3.50	< 1.50	2.50	3.50	> 3.50	< 1.50	2.50	3.50	> 3.50					
2006	0.38	0.50	0.09	0.03	0.38	0.50	0.09	0.03								0.04	0.04
2007	0.38	0.50	0.01	0.11	0.26	0.57	0.13	0.04	0.12	-0.08	-0.12	0.08				0.05	0.05
2008	0.38	0.50	0.00	0.12	0.25	0.58	0.13	0.04	0.13	-0.08	-0.13	0.08				0.05	0.05
2009	0.38	0.50	0.00	0.12	0.03	0.72	0.21	0.05	0.36	-0.22	-0.21	0.07				0.22	0.00
2009	0.04	0.50	0.32	0.15	0.04	0.50	0.32	0.15	0.00	0.00	0.00	0.00				0.00	0.00
2010	0.04	0.50	0.03	0.44	0.06	0.45	0.24	0.25	-0.02	0.04	-0.21	0.19				0.08	0.08
2011	0.04	0.50	0.00	0.47												0.39	0.17
2012	0.04	0.50	0.00	0.47													
2013	0.04	0.50	0.00	0.47													

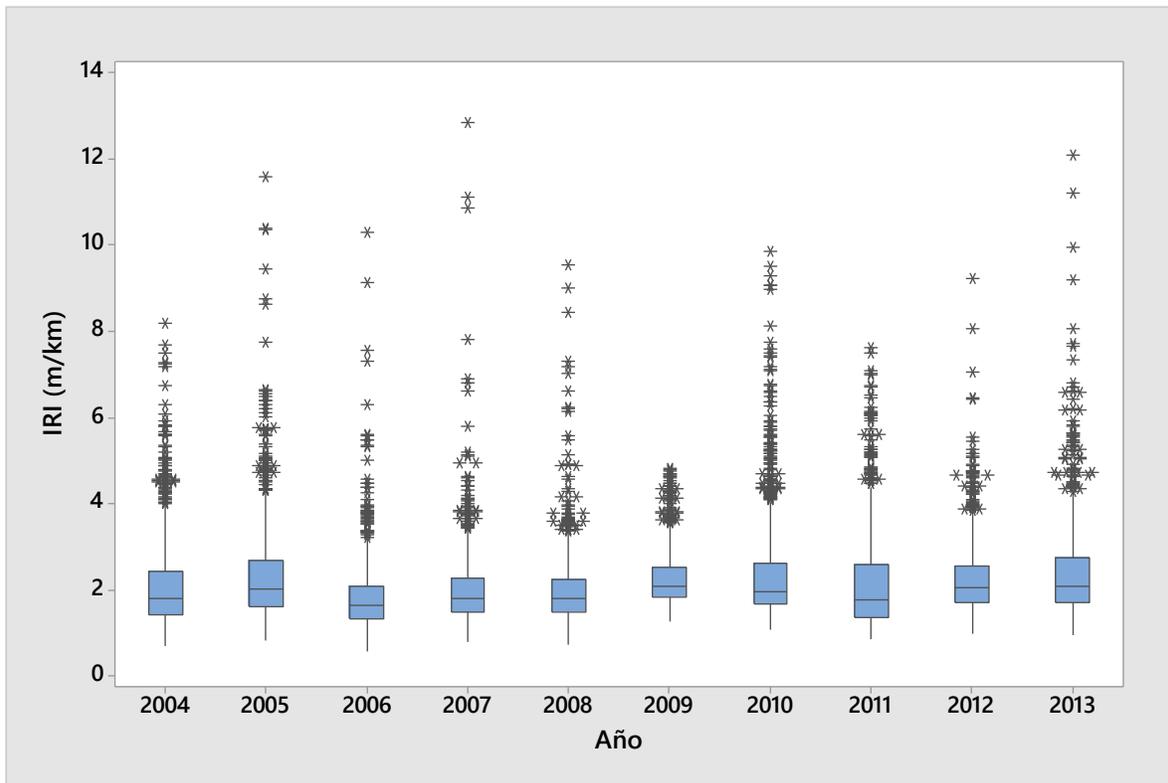


Figura 3.1 Diagrama de caja de los datos de la familia 00-ECa-DCm-PMa.

Tabla 3.5 Valores esperados de IRI en m/km, familia 00-ECa-DCm-PMa.

Año	Marcas de clase			
	< 1.50	2.50	3.50	> 3.50
2006	1.20	2.00	3.00	4.70
2009	1.41	2.00	3.00	4.10

Año	MPT promedio			
	Cadena de Márkov	Mediciones	Σerr^2_1	Σerr^2_2
2006	1.87	1.80	0.07	0.07
2007	2.27	1.99	0.28	0.28
2008	2.54	1.99	0.56	0.56
2009	2.82	2.26	0.56	0.00
2009	2.61	2.58	0.02	0.00
2010	2.84	2.94	-0.10	-0.10
2011	3.04			
2012	3.22			
2013	3.37			

0.71	0.40
------	------

Tabla 3.5 Valores esperados de IRI en m/km, fam. 00-ECa-DCm-PMa (cont.).

MPT optimizada (vectores de estado)				
Año	Cadena de Márkov	Mediciones	Σerr^2_1	Σerr^2_2
2006	1.87	1.80	0.07	0.07
2007	2.03	1.99	0.04	0.04
2008	2.17	1.99	0.19	0.19
2009	2.32	2.26	0.06	0.00
2009	2.61	2.58	0.02	0.00
2010	2.75	2.94	-0.19	-0.19
2011	2.87			
2012	2.98			
2013	3.07			

0.08	0.08
------	------

MPT optimizada (valor esperado del IRI)				
Año	Cadena de Márkov	Mediciones	Σerr^2_1	Σerr^2_2
2006	1.87	1.80	0.07	0.07
2007	2.01	1.99	0.02	0.02
2008	2.02	1.99	0.04	0.04
2009	2.03	2.26	-0.23	0.00
2009	2.61	2.58	0.02	0.00
2010	2.93	2.94	-0.01	-0.01
2011	2.96			
2012	2.96			
2013	2.96			

0.06	0.01
------	------

Cabe hacer notar que las siguientes tareas, cruciales para hacer viable la generación de modelos a partir de información histórica, no forman parte de los requerimientos antes descritos:

- Clasificación de tramos en familias y grupos, con un procedimiento similar al presentado en el inciso 1.4.
- Preparación de las series históricas que constituyen los datos de entrada del software, a partir de la información de las mediciones que se encuentra bajo resguardo de la Dirección General de Servicios Técnicos y de otras dependencias de la Secretaría.

En buena medida, este hecho es producto de que, desafortunadamente, la información original no reside en un sistema de bases de datos, sino en una colección de decenas de miles de archivos planos con formato de Excel cuyas rutas de acceso y estructura pueden variar según el año de la campaña de medición y la parte de la red federal en la que se hayan hecho las mediciones. Adicionalmente, no existe un registro de intervenciones en formato electrónico que pueda

consultarse en forma expedita para utilizar el historial de conservación de los tramos como uno de los criterios de clasificación.

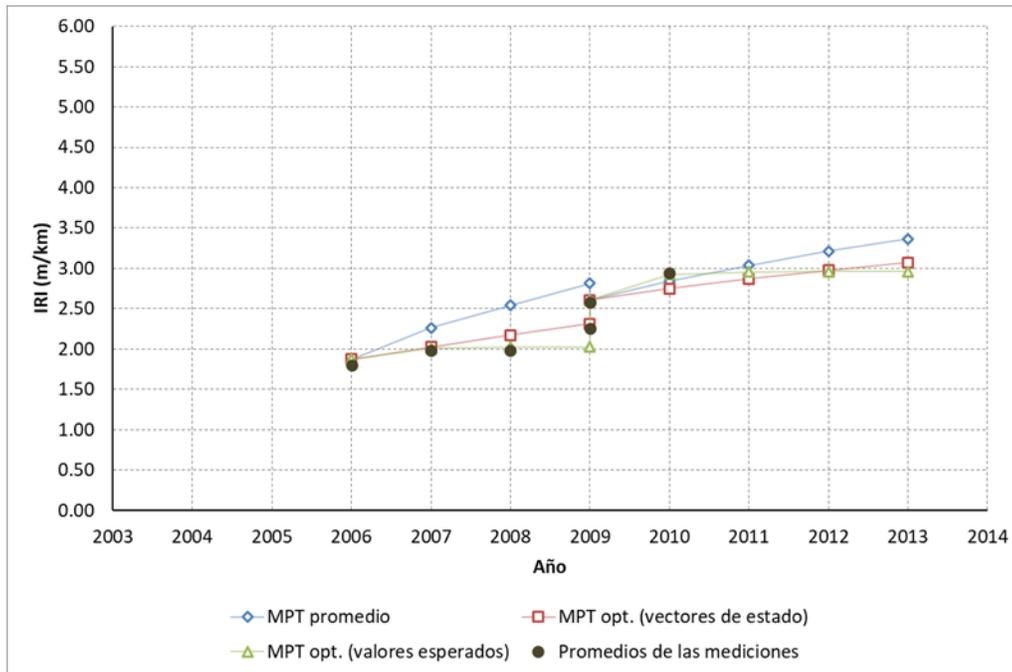


Figura 3.2 Evolución del valor esperado del IRI, familia 00-ECa-DCm-PMa.

Dado que, con las limitaciones anteriores, sería mucho más complicado automatizar esta fase inicial del proceso que cualquiera de las fases subsecuentes, se decidió no incluirla en esta primera versión del software de modelización. Así, los ejercicios de modelización que puedan hacerse en el mediano plazo estarán limitados a estudios de caso basados en la clasificación y organización manual de los datos de entrada.

3.2 Procesos principales del sistema

Los requerimientos del software abarcan algunos procesos complejos, como los siguientes:

- Lectura y preprocesado de los datos originales.
- Aplicación de las pruebas de los signos y t por pares.
- Ensamble de datos útiles para la modelización.
- Obtención de las matrices de probabilidades de transición, especialmente las que se basan en la aplicación de rutinas de optimización.
- Elaboración de gráficas a través del código de un lenguaje de programación.

Con todo, se puede apreciar que, en su conjunto, el sistema de cómputo requerido es relativamente sencillo, ya que, por una parte, se utiliza una sola entidad de datos (los datos de IRI) y se accede ella en modo de solo lectura y, por otra, los procesos que inciden sobre el flujo de información del sistema son realmente pocos, amén de

que algunos de ellos ya se encuentran implementados en varios entornos de programación.

Consecuentemente, se considera que el problema no requiere la aplicación de técnicas formales de análisis y diseño de programas para lograr una concepción adecuada del software, y que esta puede lograrse mediante una representación gráfica simplificada, como el diagrama de flujo de la Figura 3.3.

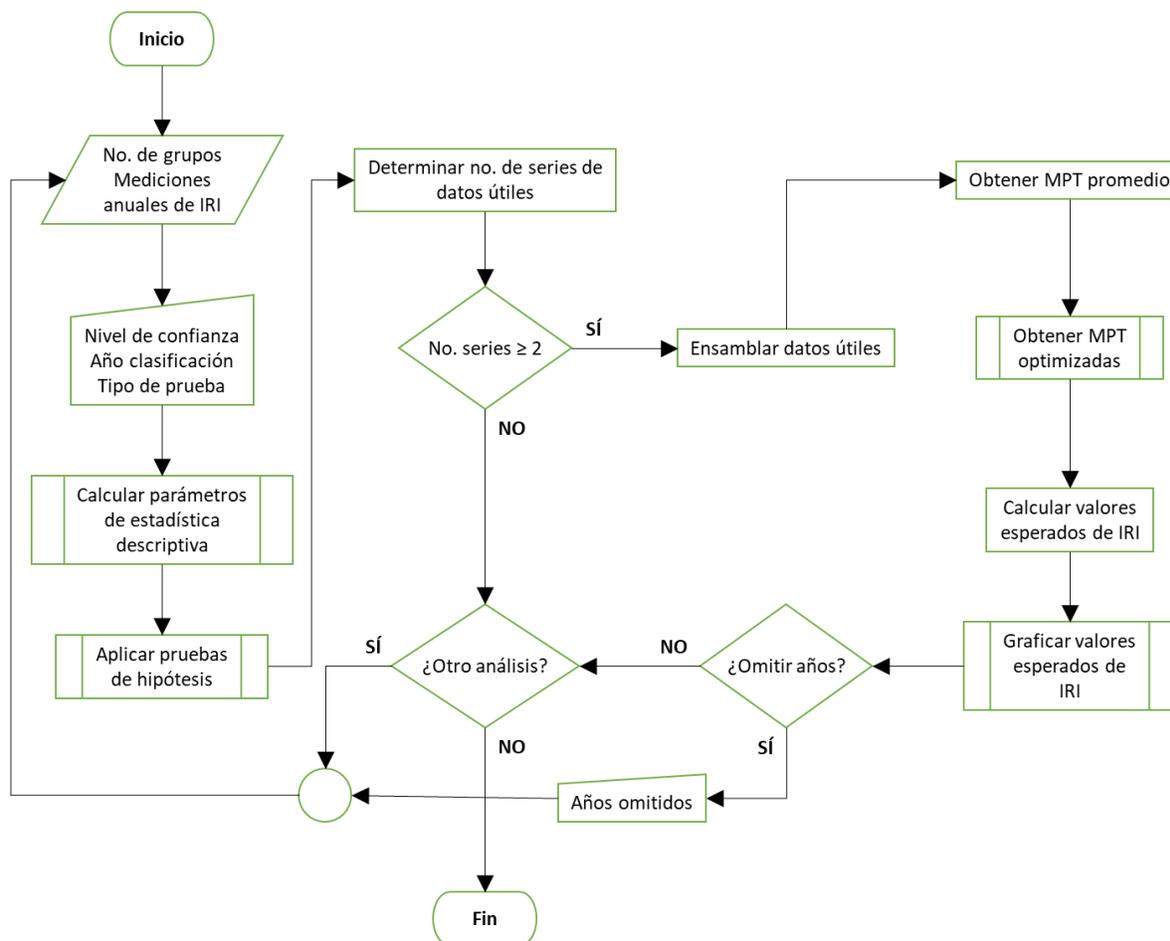


Figura 3.3 Diagrama de flujo de los procesos del software.

De este diagrama puede destacarse lo siguiente:

- Se ha utilizado el símbolo de *proceso predefinido*, es decir, el rectángulo con líneas verticales a los lados, para identificar las etapas del análisis que podrían llevarse a cabo mediante funciones o subrutinas preprogramadas en la plataforma de desarrollo elegida.
- El proceso denominado *ensamblar datos útiles* no es trivial, ya que implica agrupar en un solo juego de datos las series que corresponden a los distintos grupos de una familia, identificar las posibles discontinuidades que existan según la cobertura temporal de cada ciclo de deterioro y prever el reinicio de

las cadenas de Márkov en los años en los que se presenten las discontinuidades. Este reinicio es necesario porque en cada discontinuidad cambia el conjunto de información utilizado para definir el estado de deterioro.

- La opción *omitir años* es útil, con los datos disponibles, para ignorar la serie correspondiente a 2009 por la anomalía que se ilustra en la Figura 3.1.

En concordancia con lo antes expuesto, los procesos identificados en el diagrama de flujo, específicamente los que no se han considerado como posibles *procesos predefinidos*, se programarán como módulos separados.

3.3 Plataforma de desarrollo

Como plataforma para el desarrollo del software de modelización se ha elegido *Visual Basic for Applications* de Excel, un entorno de desarrollo integrado (*IDE* por sus siglas en inglés) incorporado a esta y a otras aplicaciones de la gama Office que ofrece un lenguaje de programación muy similar a la versión 6 de Visual Basic.

La versión de VBA de Excel es, de hecho, la más popular. Ello se debe principalmente a que, en este entorno, se pueden crear innumerables procedimientos para complementar las fórmulas y funciones de esta aplicación, las cuales, aunque variadas y en muchos casos especializadas, no siempre son suficientes para realizar ciertos cálculos y operaciones complejas.

En este caso específico, la selección de VBA de Excel se ha basado en las consideraciones que se enumeran a continuación:

- a) Los datos de entrada del proceso de modelización y los principales resultados que se desea obtener tienen un formato tabular, lo que sugiere la conveniencia de dar tratamiento a esta información en Excel como, de hecho, se hizo en la parte inicial del proyecto.
- b) Junto con el lenguaje Visual Basic, el desarrollador tiene acceso a una biblioteca muy completa de funciones matemáticas y estadísticas, entre las que pueden encontrarse las necesarias para el cálculo de parámetros de estadística descriptiva y para la estimación del valor P de las pruebas de los signos y t por pares.
- c) VBA permite emplear en el código de los programas el *algoritmo de optimización no lineal del gradiente reducido generalizado (GRG)* disponible en el paquete de herramientas *Solver* de Excel. Cabe mencionar que el algoritmo GRG de Solver se utilizó durante el desarrollo de la metodología de modelización para obtener las matrices de probabilidades de transición optimizadas en función de la minimización de los errores en las estimaciones de los vectores de estado y de los valores esperados del IRI.

- d) Adicionalmente, VBA proporciona acceso a las funciones de la biblioteca de gráficos de Office, con lo cual se simplifica notablemente la generación de salidas visuales como las curvas de evolución de los valores esperados que se especifican en la definición de requerimientos.

Para ilustrar el uso de esta plataforma, en la Figura 3.4 se muestra el editor de VBA con una parte del código empleado para el cálculo del valor P de la prueba t por pares. Nótese, en la parte inferior del editor, la llamada a la función integrada de Excel `TTest(Arr1, Arr2, Colas, Tipo_prueba)` para calcular probabilidades vinculadas a la prueba t.

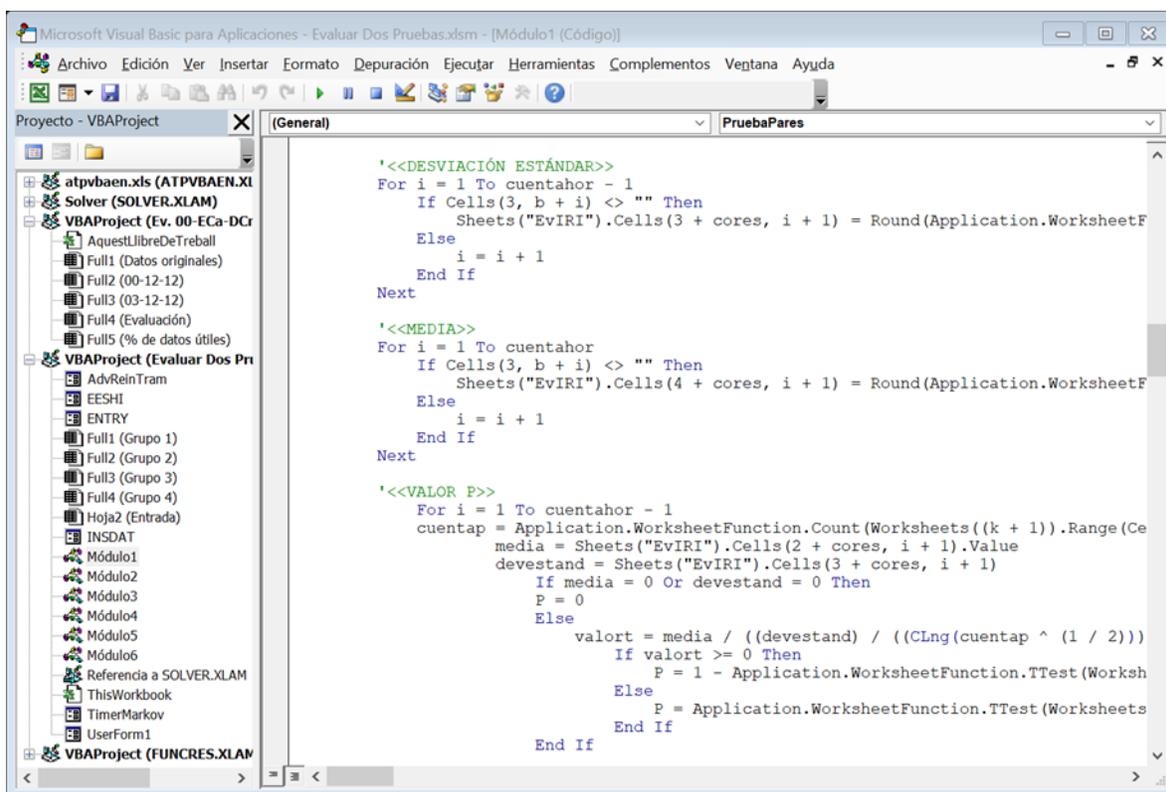


Figura 3.4 Vista parcial del código del programa en la plataforma de VBA.

3.4 Procesos estadísticos

Con referencia al diagrama de la Figura 3.3, el desarrollo del software comprende la programación de los siguientes dos procesos de índole estadística no relacionados con la generación de las cadenas de Márkov:

- Cálculo de parámetros de estadística descriptiva.
- Aplicación de pruebas de hipótesis, esto es, las pruebas de los signos y t por pares que se utilizan para delimitar los ciclos de deterioro y determinar el número de años con información útil para cada grupo de cada familia de tramos.

Estos procesos utilizan como datos de entrada las mediciones anuales de IRI de los grupos en el formato descrito como parte de la definición de requerimientos, la que además establece que los resultados deberán presentarse en un formato similar al de la Tabla 3.1.

En la misma definición, se indica que deben calcularse los siguientes parámetros estadísticos para las diferencias entre mediciones anuales consecutivas: media, desviación estándar y mediana. Para este fin, se utilizaron las respectivas funciones incorporadas de Excel, es decir:

```
Average (Arg1, ..., Arg30)  
StdDev (Arg1, ..., Arg30)  
Median (Arg1, ..., Arg30)
```

en las que los argumentos `Arg1, ..., Arg30` pueden representar datos individuales, nombres, arreglos o referencias que contienen nombres. En el caso del software se utilizaron arreglos con las diferencias entre las mediciones. Evidentemente, los cálculos respectivos pudieron haberse codificado directamente en VBA, sin embargo, se observó que el uso de las funciones incorporadas se traduce en un decremento de los tiempos de ejecución y procesamiento de los datos, además de que también reduce el tiempo invertido en la programación.

Las funciones antes citadas también se utilizan para el cálculo de los porcentajes de cambio de la media y la desviación estándar entre las series consecutivas.

Por lo que respecta a la aplicación de las pruebas de hipótesis y, específicamente, al cálculo del valor P, en el caso de la prueba de los signos se utilizó la función incorporada de distribución binomial negativa:

```
NegBinom_Dist (Núm_fracasos, Núm_éxitos, Prob_éxito, Acum)
```

En la expresión anterior:

`Num_fracasos`: Número de fracasos.

`Núm_éxitos`: Número límite de éxitos.

`Prob_éxito`: Probabilidad de obtener un éxito.

`Acum`: Valor lógico que determina la modalidad de la función. `VERDADERO` refiere a la distribución acumulada y `FALSO` a la función de densidad de probabilidad.

Para la aplicación al cálculo del valor P de la prueba de los signos, se utilizan, en orden respectivo, los siguientes valores de los argumentos: número de diferencias entre mediciones mayores a cero, tamaño de la muestra, 0.5 y "VERDADERO".

La estimación del valor P que corresponde a la prueba t por pares se basa en la función incorporada en Excel para la misma prueba t, es decir:

`TTest(Arr1, Arr2, Colas, Tipo_prueba)`

donde:

`Arr1`: Primer conjunto de datos.

`Arr2`: Segundo conjunto de datos.

`Colas`: Número de colas de la distribución.

`Tipo_prueba`: Tipo de prueba.

En el contexto de aplicación del software, `Arr1` y `Arr2` representan las dos series consecutivas, y se utiliza una prueba t de una sola cola (`Colas=1`) y por pares (`Tipo_prueba=1`).

Adicionalmente, se hace la siguiente consideración para el cálculo del valor P:

$$t_0 < 0 \Rightarrow P = T.Test(Serie1, Serie2, 1, 1)$$

$$t_0 > 0 \Rightarrow P = 1 - T.Test(Serie1, Serie2, 1, 1)$$

Para finalizar el proceso de análisis estadístico, el software identifica los ciclos de deterioro útiles de los grupos de cada familia, que consisten en periodos de dos años o más en los que la aplicación de las pruebas de hipótesis produce un resultado de $P \geq 0.05$ y, al mismo tiempo, incluye el año en el que fueron recabados los datos en los que se basa la clasificación de los tramos. Como se recuerda, en el caso de los datos usados en el proyecto este año es 2008.

El software permite al usuario especificar el nivel de confianza que desea utilizar en el análisis, así como realizar un análisis por familia de tramos (incluyendo todos sus grupos) a la vez.

En la Tabla 3.6 se presentan los resultados de la prueba de los signos obtenidos con el software para la familia *00-ECa-DCm-PMa*. Además de reproducir el formato especificado en la definición de requerimientos, se incluyen celdas con los periodos de años útiles determinados para cada grupo de la familia. Si alguna de estas celdas colocadas en forma secuencial hacia abajo aparece vacía, debe asumirse que no se encontraron datos útiles para el correspondiente grupo.

Por otro lado, la Tabla 3.7 contiene la estimación realizada para la misma familia mediante la prueba t por pares. Como se observa, para esta prueba no se incluyen resultados relativos a la mediana, ya que este parámetro es irrelevante para la prueba t por lo que se refiere al contraste de las hipótesis.

Tabla 3.6 Resultados de la prueba de los signos, familia 00-ECa-DCm-PMa.

Años	Años útiles												
	2004 - 2005	2005 - 2006	2006 - 2007	2007 - 2008	2008 - 2009	2009 - 2010	2010 - 2011	2011 - 2012	2012 - 2013	(2006-2010)	(2006-2009)		
Media	-0.071	-0.756	0.089	-0.054	0.111	0.357	-0.381	-0.270	0.947				
Desv. est.	1.694	1.636	1.856	1.102	1.492	1.203	1.475	1.597	1.769				
Mediana	-0.095	-0.842	0.117	-0.062	0.200	0.062	-0.247	-0.085	0.775				
Valor P	0.285	0.000	0.927	0.057	0.990	0.886	0.000	0.195	1.000				
Inc. Media	-2.29%	-23.38%	3.99%	-2.18%	4.39%	13.88%	-11.33%	-14.32%	48.40%				
Inc. desv. est.	8.07%	-2.25%	12.26%	-14.80%	-29.61%	90.69%	-12.15%	-36.36%	74.49%				

Años	Años útiles												
	2004 - 2005	2005 - 2006	2006 - 2007	2007 - 2008	2008 - 2009	2009 - 2010	2010 - 2011	2011 - 2012	2012 - 2013	(2006-2010)	(2006-2009)		
Media	0.304	-0.439	0.214	0.014	0.309	0.062	-0.202	0.175	-0.026				
Desv. est.	1.088	1.009	0.759	0.746	0.788	0.868	1.097	1.170	1.131				
Mediana	0.250	-0.345	0.175	0.010	0.325	-0.145	-0.210	0.285	-0.040				
Valor P	1.000	0.000	1.000	0.626	1.000	0.000	0.000	1.000	0.089				
Inc. Media	17.02%	-21.08%	12.92%	0.37%	17.22%	2.16%	-8.07%	8.90%	-1.15%				
Inc. desv. est.	45.61%	-42.50%	13.99%	6.62%	-19.07%	89.15%	1.33%	-18.16%	4.35%				

Tabla 3.7 Resultados de la prueba t por pares, familia 00-ECa-DCm-PMa.

Años	Años útiles												
	2004 - 2005	2005 - 2006	2006 - 2007	2007 - 2008	2008 - 2009	2009 - 2010	2010 - 2011	2011 - 2012	2012 - 2013	(2006-2010)	(2006-2009)		
Media	-0.071	-0.756	0.089	-0.054	0.111	0.357	-0.381	-0.270	0.947				
Desv. est.	1.694	1.636	1.856	1.102	1.492	1.203	1.475	1.597	1.769				
Valor P	0.254	0.000	0.776	0.218	0.880	1.000	0.000	0.001	1.000				
Inc. Media	-0.071	-0.756	0.089	-0.054	0.111	0.357	-0.381	-0.270	0.947				
Inc. desv. est.	1.694	1.636	1.856	1.102	1.492	1.203	1.475	1.597	1.769				

Años	Años útiles												
	2004 - 2005	2005 - 2006	2006 - 2007	2007 - 2008	2008 - 2009	2009 - 2010	2010 - 2011	2011 - 2012	2012 - 2013	(2006-2010)	(2006-2009)		
Media	0.304	-0.439	0.214	0.014	0.309	0.062	-0.202	0.175	-0.026				
Desv. est.	1.088	1.009	0.759	0.746	0.788	0.868	1.097	1.170	1.131				
Valor P	1.000	0.000	1.000	0.727	1.000	0.987	0.000	1.000	0.238				
Inc. Media	17.02%	-21.08%	12.92%	0.37%	17.22%	2.16%	-8.07%	8.90%	-1.15%				
Inc. desv. est.	45.61%	-42.50%	13.99%	6.62%	-19.07%	89.15%	1.33%	-18.16%	4.35%				

Con referencia a las Tablas 3.1 y 3.6, los resultados asentados en ambas son prácticamente idénticos, lo que demuestra la operatividad del software con respecto al análisis estadístico. Las pocas y pequeñas diferencias que se observan pueden ser atribuibles a los métodos empleados en cada caso, a errores de aproximación de Excel o a errores al registrar los resultados en el primer análisis, el cual fue realizado en forma manual.

En lo que concierne a la prueba t por pares, dada la mayor potencia de esta prueba, cabía esperar una menor extensión de los ciclos de deterioro. Sin embargo, en el caso representado, los ciclos resultaron con el mismo número de años útiles para el primer grupo, y con un año adicional para el segundo. No obstante, como puede verificarse en los resultados de la comparación global de las pruebas que se presenta en el capítulo 4, en general la prueba t por produjo ciclos más cortos para la mitad de las familias, lo que, entre otros efectos, produjo una notable reducción del error vinculado a las estimaciones de los modelos. Esta reducción, podría ser un reflejo de la sospechada ocurrencia de errores tipo II con la prueba de los signos.

3.5 Obtención de las cadenas de Márkov

De acuerdo con lo estipulado en la fase de definición de requerimientos, el software debe proveer las funciones necesarias para implementar los siguientes procesos relacionados con la generación de las cadenas de Márkov:

- a) Obtener las matrices de probabilidades de transición (MPT) con los siguientes tres métodos:
 - Cálculo de los promedios de las proporciones de tramos que pasan de un estado de deterioro a otro durante el periodo definido por el ciclo de deterioro (ec. 1.6).
 - Optimización basada en la minimización de las diferencias entre los vectores de estado pronosticados con la MPT y los calculados con las series originales de datos (ec. 1.7).
 - Optimización basada en la minimización de las diferencias entre los valores esperados de IRI estimados a partir de las cadenas de Márkov y los reflejados por las series originales (ec. 1.8).
- b) Generar cadenas de Márkov con cada una de las MPT obtenidas (ec. 1.5) y compararlas con las que pueden extraerse de los datos originales, en términos de la suma de los cuadrados de las diferencias.
- c) Calcular los valores esperados de IRI con cada una de las cadenas generadas (ecs. 1.10 y 1.11) y compararlos con los promedios de las mediciones anuales, expresando nuevamente los errores como sumas de cuadrados.
- d) Graficar la evolución del IRI en el tiempo con los valores esperados del IRI del punto anterior.

Como punto de partida para la aplicación de los procesos anteriores, y de acuerdo con la Figura 3.3, se escribió el código necesario para ensamblar el conjunto de datos útiles, esto es, la fracción de los datos originales de cada familia que se ubica dentro de los periodos de aceptación determinados para los diferentes grupos con el auxilio de las pruebas estadísticas. Asimismo, se hicieron las previsiones necesarias para identificar las discontinuidades producidas por la falta de coincidencia entre los periodos de los grupos, la cual debe tomarse en cuenta en la generación de las cadenas de Márkov.

En relación con la matriz basada en los promedios de las proporciones o *matriz promedio*, dado que el método de cálculo respectivo utiliza directamente las mediciones consecutivas de IRI, antes de aplicar este método es necesario eliminar las diferencias negativas que existan, ya que de no hacerlo podrían aparecer elementos en la MPT con este mismo signo. Para ello, el software aplica un filtro diseñado en el estudio original que iguala con cero las diferencias negativas de las parejas cuyos valores de IRI se encuentran en el mismo rango (ver Tabla 3.2) y descarta las parejas con datos en rangos distintos. Adicionalmente, el filtro limita la magnitud de las diferencias positivas de manera que se cumpla el supuesto representado en la Figura 1.4 (b) de que un pavimento no puede deteriorarse en más de un estado durante un año.

Por lo que toca a la obtención de las MPT que involucran procesos de optimización, el software aplica el algoritmo GRG de Solver (Fylstra, et al., 1998). En la Figura 3.5 se muestra un fragmento del código del programa que se relaciona con la aplicación de este algoritmo.

```
'SOLVER OPTIMIZACION 1
  Sheets("Optimizacion 1").Activate
  Sheets("Optimizacion 1").Select
  SolverReset
  SolverOptions precision:=0.0001
  SolverOk SetCell:="$H$2", MaxMinVal:=2,
    ValueOf:=0, ByChange:="$B$3:$C$3,$C$4:$D$4,$D$5:$E$5", _
    Engine:=1, EngineDesc:="GRG Nonlinear"
  SolverAdd cellRef:="$B$3:$C$3", relation:=1, formulaText:="1"
  SolverAdd cellRef:="$B$3:$C$3", relation:=3, formulaText:="0"
  SolverAdd cellRef:="$C$4:$D$4", relation:=1, formulaText:="1"
  SolverAdd cellRef:="$C$4:$D$4", relation:=3, formulaText:="0"
  SolverAdd cellRef:="$D$5:$E$5", relation:=1, formulaText:="1"
  SolverAdd cellRef:="$D$5:$E$5", relation:=3, formulaText:="0"
  SolverAdd cellRef:="$F$3:$F$6", relation:=2, formulaText:="1"
  SolverSolve userFinish:=True
```

Figura 3.5 Fragmento de código relacionado con el algoritmo GRG.

Una vez calculadas las diferentes versiones de las MPT, exceptuando la gráfica de los valores estimados, el resto de los productos de cadenas de Márkov pueden obtenerse mediante operaciones matriciales, vectoriales o aritméticas relativamente sencillas. En cuanto a la gráfica, para generarla desde el código y como se había previsto, se utilizó la biblioteca de objetos gráficos de Office a través de VBA.

En las tablas 3.8 a 3.10 se presentan las MPT, cadenas de Márkov y valores esperados de IRI generados con el software para la familia *00-ECa-DCm-PMa*. La Figura 3.6 muestra en forma gráfica la evolución del IRI esperado durante el ciclo de deterioro. Todos estos productos provienen de los datos útiles delimitados con la prueba de los signos.

Si se comparan los elementos anteriores con sus contrapartes de la definición de requerimientos, se puede verificar que, en todos los casos, se obtuvieron resultados prácticamente idénticos. De este modo, queda también demostrada la operatividad del software en lo que se refiere al desarrollo de modelos de deterioro basados en cadenas de Márkov.

Los modelos de la familia *00-ECa-DCm-PMa* también se desarrollaron con los datos útiles identificados con la prueba t por pares. Aunque igualmente se obtuvieron los resultados tabulares, en esta sección solo se reproduce el gráfico de evolución de los valores esperados de IRI (Figura 3.7), a fin de compararlo con el que corresponde a la prueba de los signos.

Las Tablas 3.6 y 3.7 muestran que la principal diferencia entre los ciclos de deterioro identificados por cada prueba para la familia *00-ECa-DCm-PMa*, es que, mientras que para la prueba de los signos el ciclo del grupo 2 concluye en 2009, en el caso de la prueba t por pares este ciclo se extiende hasta el año 2010.

El resultado de la prueba de los signos crea una discontinuidad en 2009 en el paquete de datos útiles, la cual se traduce en un salto en el gráfico de evolución de los valores esperados de la Figura 3.6 por el “reinicio” de la cadena de Márkov en ese año. En el análisis con la prueba t por pares los dos ciclos de deterioro resultaron con la misma extensión, de manera que la Figura 3.7 no presenta el salto antes referido y los modelos obtenidos con la segunda prueba generan un menor error, por lo menos en lo que corresponde a las curvas calculadas con las MPT optimizadas.

Tabla 3.8 Obtención de las MPT, familia *00-ECa-DCm-PMa*.

		MPT promedio			
		< 1.50	2.50	3.50	> 3.50
1.50		0.20	0.80	0.00	0.00
2.50		0.00	0.80	0.20	0.00
3.50		0.00	0.00	0.68	0.32
> 3.50		0.00	0.00	0.00	1.00

		MPT opt. (vectores de estado)						MPT opt. (valores esperados)			
		< 1.50	2.50	3.50	> 3.50			< 1.50	2.50	3.50	> 3.50
1.50		0.76	0.24	0.00	0.00	1.50		1.00	0.00	0.00	0.00
2.50		0.00	0.92	0.08	0.00	2.50		0.00	1.00	0.00	0.00
3.50		0.00	0.00	0.73	0.27	3.50		0.00	0.00	0.09	0.91
> 3.50		0.00	0.00	0.00	1.00	> 3.50		0.00	0.00	0.00	1.00

Tabla 3.9 Desarrollo de las cadenas de Márkov, familia 00-ECa-DCm-PMa.

Año	MPT promedio														
	Vectores de estado calculados			Vectores de estado originales			Diferencias			Σerr^2_1 Σerr^2_2					
	< 1.50	2.50	3.50	< 1.50	2.50	3.50	< 1.50	2.50	3.50	< 1.50	2.50	3.50	> 3.50	Σerr^2_1	Σerr^2_2
2006	0.38	0.50	0.09	0.03	0.38	0.50	0.09	0.03	0.00	0.00	0.00	0.00	0.00	0.05	0.05
2007	0.08	0.70	0.16	0.06	0.26	0.57	0.13	0.04	-0.18	0.12	0.03	0.02	0.07	0.08	0.08
2008	0.02	0.62	0.25	0.11	0.25	0.58	0.13	0.04	-0.23	0.04	0.12	0.07	0.07	0.08	0.08
2009	0.00	0.51	0.30	0.19	0.03	0.71	0.21	0.05	-0.02	-0.21	0.09	0.14	0.07	0.07	0.00
2009	0.04	0.50	0.32	0.15	0.04	0.50	0.32	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2010	0.01	0.42	0.32	0.25	0.06	0.45	0.24	0.25	-0.05	-0.02	0.08	0.00	0.00	0.01	0.01
2011	0.00	0.34	0.30	0.35										0.21	0.14
2012	0.00	0.28	0.28	0.45											
2013	0.00	0.22	0.25	0.53											

Año	MPT optimizada (vectores de estado)														
	Vectores de estado calculados			Vectores de estado originales			Diferencias			Σerr^2_1 Σerr^2_2					
	< 1.50	2.50	3.50	< 1.50	2.50	3.50	< 1.50	2.50	3.50	< 1.50	2.50	3.50	> 3.50	Σerr^2_1	Σerr^2_2
2006	0.38	0.50	0.09	0.03	0.38	0.50	0.09	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2007	0.29	0.55	0.10	0.06	0.26	0.57	0.13	0.04	0.03	-0.02	-0.02	0.02	0.02	0.00	0.00
2008	0.22	0.58	0.12	0.09	0.25	0.58	0.13	0.04	-0.03	0.00	-0.02	0.05	0.05	0.00	0.00
2009	0.16	0.59	0.13	0.12	0.03	0.71	0.21	0.05	0.14	-0.13	-0.08	0.07	0.07	0.05	0.00
2009	0.04	0.50	0.32	0.15	0.04	0.50	0.32	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2010	0.03	0.47	0.27	0.24	0.06	0.45	0.24	0.25	-0.03	0.02	0.03	-0.01	0.00	0.00	0.00
2011	0.02	0.44	0.23	0.31										0.05	0.01
2012	0.02	0.41	0.20	0.37											
2013	0.01	0.38	0.18	0.43											

Tabla 3.7 Desarrollo de las cadenas de Márkov, familia 00-ECa-DCm-PMa (cont.).

Año	MPT optimizada (valores esperados)										Σerr^2_1	Σerr^2_2		
	Vectores de estado calculados		Vectores de estado originales				Diferencias							
	< 1.50	2.50	3.50	> 3.50	< 1.50	2.50	3.50	> 3.50	< 1.50	2.50			3.50	> 3.50
2006	0.38	0.50	0.09	0.03	0.38	0.50	0.09	0.03	0.00	0.00	0.00	0.00	0.04	0.04
2007	0.38	0.50	0.01	0.11	0.26	0.57	0.13	0.04	0.12	-0.08	-0.12	0.08	0.05	0.05
2008	0.38	0.50	0.00	0.12	0.25	0.58	0.13	0.04	0.13	-0.08	-0.13	0.08	0.05	0.05
2009	0.38	0.50	0.00	0.12	0.03	0.71	0.21	0.05	0.36	-0.22	-0.21	0.07	0.22	0.00
2009	0.04	0.50	0.32	0.15	0.04	0.50	0.32	0.15	0.00	0.00	0.00	0.00	0.00	0.00
2010	0.04	0.50	0.03	0.44	0.06	0.45	0.24	0.25	-0.02	0.05	-0.21	0.19	0.08	0.08
2011	0.04	0.50	0.00	0.47									0.39	0.17
2012	0.04	0.50	0.00	0.47										
2013	0.04	0.50	0.00	0.47										

Tabla 3.10 Cálculo de valores esperados, familia 00-ECa-DCm-PMa.

Año	Marcas de clase			
	< 1.50	2.50	3.50	> 3.50
2006	1.20	2.00	3.00	4.70
2009	1.41	2.00	3.00	4.10

Año	Cadena de Márkov	Mediciones	MPT promedio	
			Σerr^2_1	Σerr^2_2
2006	1.87	1.80	0.07	0.07
2007	2.27	1.99	0.28	0.28
2008	2.54	1.99	0.56	0.56
2009	2.82	2.26	0.56	0.00
2009	2.69	2.58	0.11	0.00
2010	2.84	2.94	-0.10	-0.10
2011	3.04			
2012	3.22			
2013	3.37			
			0.73	0.40

Año	Cadena de Márkov	Mediciones	MPT optimizada (vectores de estado)	
			Σerr^2_1	Σerr^2_2
2006	1.87	1.80	0.07	0.07
2007	2.03	1.99	0.04	0.04
2008	2.17	1.99	0.18	0.18
2009	2.28	2.26	0.02	0.00
2009	2.61	2.58	0.03	0.00
2010	2.75	2.94	-0.19	-0.19
2011	2.87			
2012	2.98			
2013	3.07			
			0.08	0.07

Año	Cadena de Márkov	Mediciones	MPT optimizada (valor esperado del IRI)	
			Σerr^2_1	Σerr^2_2
2006	1.87	1.80	0.07	0.07
2007	2.01	1.99	0.02	0.02
2008	2.02	1.99	0.03	0.03
2009	2.03	2.26	-0.23	0.00
2009	2.61	2.58	0.35	0.00
2010	2.93	2.94	-0.01	-0.01
2011	2.96			
2012	2.96			
2013	2.96			
			0.06	0.01

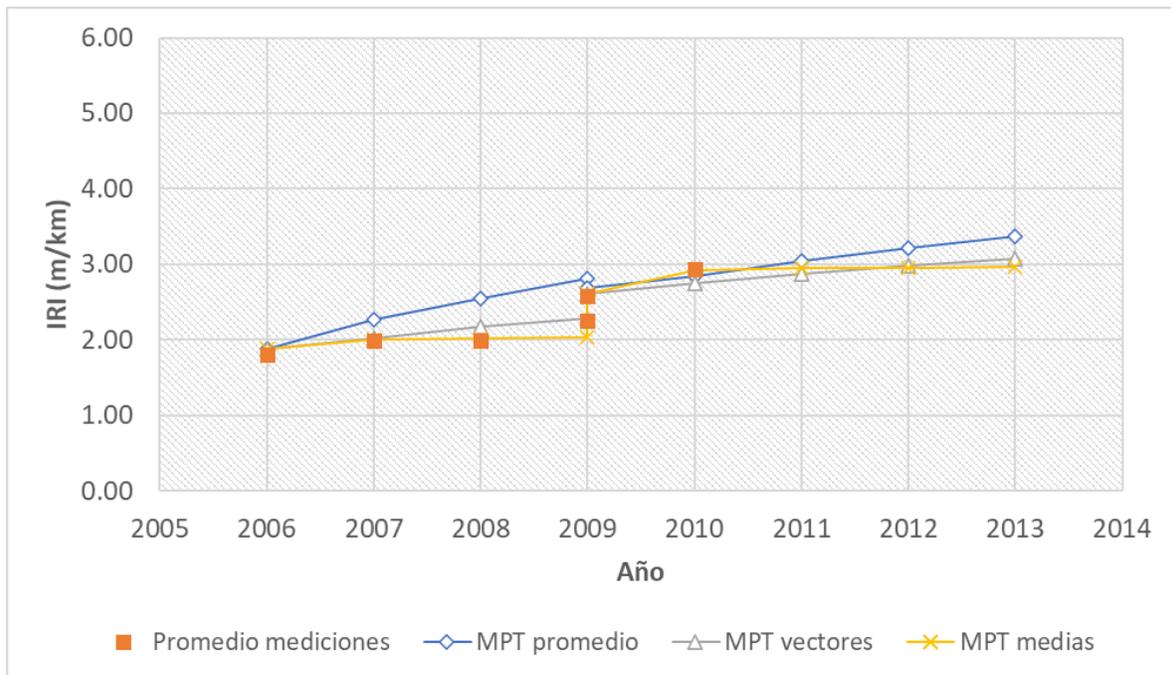


Figura 3.6 Modelos de la familia 00-ECa-DCm-PMa, prueba de los signos.

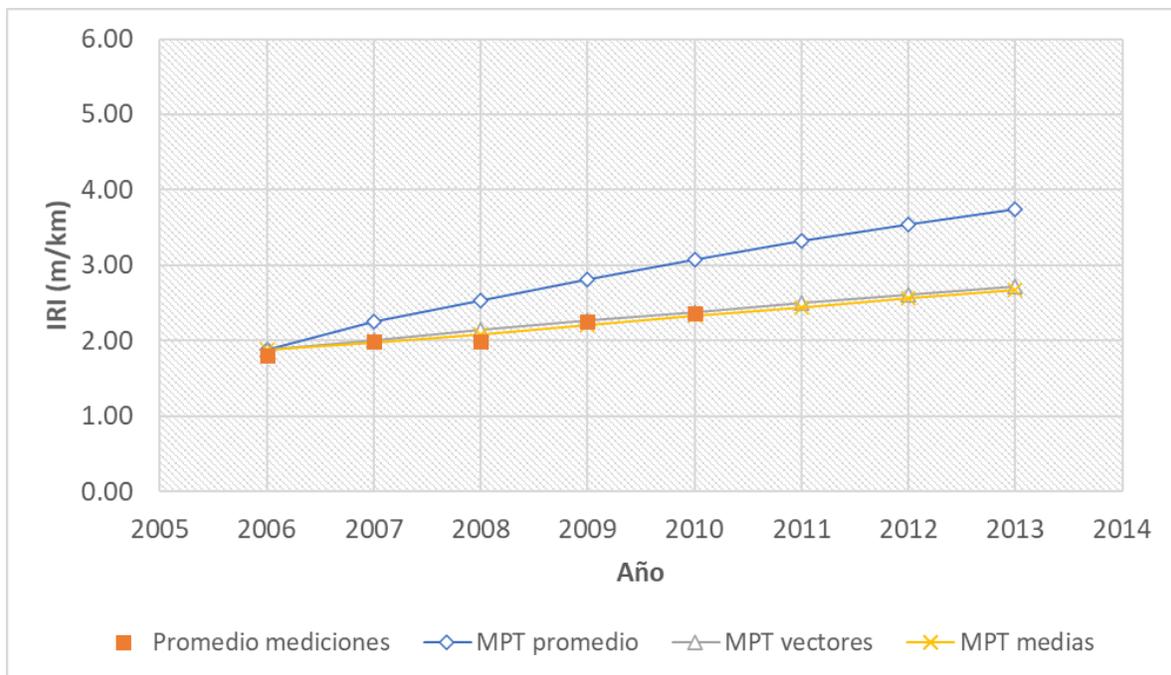


Figura 3.7 Modelos de la familia 00-ECa-DCm-PMa, prueba t por pares.

Por lo que se refiere a las curvas vinculadas a la MPT promedio, en ambos casos puede observarse que ambas despegan en un grado mayor de los datos reales que las curvas de las matrices optimizadas. Una posible explicación de este hecho, es que la aplicación del filtro para eliminar las diferencias negativas en el primer caso podría estar induciendo incrementos artificiales en la modelización basada en las MPT promedio.

3.6 Interfaz de usuario

La interfaz de usuario es el mecanismo a través del cual el usuario interactúa con el software, a fin de ingresar, por medios diversos, los datos necesarios para la aplicación de los diferentes procesos previstos en la concepción de la herramienta, así como para recuperar los resultados esperados. En términos generales, se espera que cualquier interfaz de usuario cumpla con los siguientes principios básicos:

- *Familiaridad.* Utilizar objetos y términos que resulten familiares para los usuarios.
- *Congruencia.* Conservar el formato y significado de los diferentes objetos (menús, cuadros de diálogo, botones y mensajes) en toda la aplicación.
- *Sorpresa mínima.* Asegurar que las mismas acciones en contextos comparables produzcan efectos comparables.
- *Capacidad de recuperación.* Hacer las provisiones necesarias para lograr la recuperación del programa frente a errores cometidos por el usuario o frente a otros eventos inesperados.
- *Flexibilidad.* Considerar el comportamiento diverso que corresponde a la naturaleza también diversa de los usuarios.
- *Disponibilidad de una guía.* Contemplar la preparación de una guía que sirva al usuario como referencia básica para comprender el funcionamiento del programa o para la solución de problemas.

El software se ha desarrollado sobre un entorno de programación integrado que funciona en el marco de Excel, es decir, una aplicación preexistente. Por lo tanto, es claro que comparte una buena parte de los elementos de la interfaz de usuario de esta aplicación, aunque también incorpora algunos elementos propios.

Cuando se inicia el software de modelización, el cual aún se conoce con el nombre de trabajo “Evaluación estadística de series históricas del IRI” (siglas “EESHI”), se muestra la pantalla de bienvenida reproducida en la Figura 3.8, seguida por el menú principal de la herramienta que se presenta en la Figura 3.9. Este menú está montado sobre la primera hoja de un libro de Excel (el libro principal de la aplicación), y contiene una serie de opciones cuyas funciones y orden en el que deben elegirse se indican a continuación:

- a) *Cerrar libros.* Cierra libros de Excel no relacionados con el software. En realidad, puede hacerse uso de ella en cualquier momento.

- b) *Insertar grupos*. Permite insertar las series de IRI de uno o más de los grupos de una familia mediante operaciones de “copiar y pegar”. Por cada grupo insertado, el software agrega una hoja de cálculo al libro principal.
- c) *Insertar familia*. Opción alternativa a la anterior mediante la cual se pueden agregar las series de todos los grupos de una familia, almacenadas en un archivo con el formato al que se hace referencia en la definición de requerimientos. El software lee el archivo mediante una operación de “Abrir archivo” similar a la de cualquier aplicación de Windows, y añade al libro principal tantas hojas como grupos formen parte de la familia.



Figura 3.8 Pantalla de bienvenida del software.



Figura 3.9 Menú principal.

- d) *Evaluar series*. Aplica la prueba de los signos o la prueba t por pares a los datos de una familia y determina los ciclos con datos útiles para la modelización. Requiere como datos adicionales el nivel de confianza, el año al que corresponden los datos utilizados para la clasificación de los tramos y el tipo de prueba de hipótesis que habrá de aplicarse. (Figura 3.10).



Figura 3.10 Diálogo para la definición de parámetros estadísticos.

- e) *Obtener modelos*. Genera las cadenas de Márkov y demás resultados de modelización para el paquete de datos útiles delimitado según la prueba de hipótesis.
- f) *Omitir años*. Permite al usuario omitir los datos de alguno de los años comprendidos en el ciclo de deterioro.
- g) *Finalizar*. Termina la aplicación.

Es importante destacar que, cada vez que se realiza la evaluación de un conjunto de series o que se generan modelos, el software crea libros de Excel independientes en los que vacía los resultados con los formatos especificados, para cada caso, en la definición de requerimientos, los cuales pueden ser almacenados por el usuario de acuerdo con sus necesidades.

4. Comparación de las pruebas

Como se ha mencionado previamente, uno de los objetivos de este trabajo consistió en verificar la aplicabilidad de la prueba t por pares en la delimitación de los ciclos de deterioro, dados el requisito de esta prueba en el sentido de que las diferencias de los pares se ajusten a una distribución normal y el hecho de que las diferencias de mediciones consecutivas de IRI no cumplen con este requisito.

Sin embargo, en el inciso 2.2 quedó establecido que la aplicación de la prueba t por pares puede ser válida si el conjunto de datos de interés es continuo, unimodal y “razonablemente simétrico”, condiciones que en lo general sí cumplen los datos de IRI.

La motivación por explorar la prueba t por pares surge de la sospecha de que la prueba no paramétrica de los signos, utilizada en primera instancia para la delimitación de los ciclos, podría estar generando errores tipo II.

Tomando en cuenta los antecedentes anteriores, se hicieron corridas completas del software para las trece familias de tramos con las dos pruebas, a efecto de comparar la extensión de los ciclos de deterioro producidos por cada prueba, la magnitud de los errores inducidos por los modelos y la forma de las curvas de deterioro.

En el resto del capítulo se presenta un resumen de los hallazgos de este ejercicio de comparación.

4.1 Años útiles y errores por prueba

La Tabla 4.1 muestra resultados generales de la modelización del IRI para las 13 familias, los cuales se obtuvieron mediante la aplicación de las pruebas de los signos y t por pares para delimitar los ciclos con datos útiles. Esta tabla indica, por un lado, la extensión del ciclo determinada en cada caso y, por otro, el error en la estimación de la evolución de los vectores de estado a partir de la optimización de la MPT basada en esos mismos objetos. La magnitud del error se representa gráficamente empleando una escala de colores que va del rojo (error más alto) al verde (el más bajo).

La inspección de la Tabla 4.1 permite llegar a las siguientes conclusiones:

- a) De los diez años originalmente cubiertos por el paquete de datos de entrada disponible, la prueba de los signos reduce la duración de los ciclos con datos útiles a 5.2 años en promedio y la t por pares a 4.2.

Tabla 4.1 Años útiles y errores según el tipo de prueba.

No.	Familia	Prueba de los signos		Prueba t por pares	
		Años útiles	$\Sigma \epsilon^2$	Años útiles	$\Sigma \epsilon^2$
1	00-ECa-DCb-PMm	2004 - 2010	0.045	2004 - 2009	0.036
2	00-ECa-DCm-PMa	2006 - 2010	0.008	2006 - 2010	0.004
3	00-ECa-DCm-PMm	2004 - 2008	0.066	2004 - 2008	0.066
4	00-ECb-DCb-PMa	2004 - 2008	0.101	2005 - 2008	0.029
5	00-ECb-DCb-PMm	2005 - 2010	0.060	2005 - 2008	0.041
6	00-ECb-DCm-PMa	2004 - 2010	0.062	2004 - 2008	0.035
7	00-ECb-DCm-PMm	2005 - 2010	0.062	2005 - 2008	0.009
8	00-ECm-DCa-PMa	2004 - 2008	0.029	2004 - 2008	0.029
9	00-ECm-DCb-PMa	2004 - 2010	0.021	2004 - 2008	0.008
10	00-ECm-DCb-PMm	2004 - 2012	0.105	2006 - 2012	0.070
11	00-ECm-DCm-PMa	2004 - 2011	0.062	2004 - 2011	0.074
12	00-ECm-DCm-PMm	2004 - 2008	0.014	2004 - 2008	0.014
13	20-ECa-DCb-PMa	2006 - 2010	0.019	2006 - 2010	0.019
		Suma:	0.656	Suma:	0.434

Reducción del error: 34%

- b) La prueba t por pares hace más corta la duración con respecto a la de la prueba de los signos para 7 de las 13 familias, la deja igual para seis y no la aumenta en ningún caso.
- c) La menor extensión de los ciclos de deterioro con respecto al periodo que abarcan los datos responde a la detección de series consecutivas con diferencias negativas estadísticamente significativas, que no son compatibles con la modelización del deterioro.
- d) La extensión aún menor que se obtiene en algunos casos con la prueba t por pares, sugiere que esta prueba tiene mayor capacidad para detectar evidencia de tendencias negativas y, correspondientemente, que podrían estarse cometiendo errores tipo II con la prueba de los signos.
- e) Los modelos basados en datos evaluados por la prueba t tuvieron un error menor que los de la prueba de los signos en 8 de los casos, igual en 4 y mayor en 1.
- f) El error total de las estimaciones es 34 % menor en el caso de la prueba t por pares. En general, los menores errores vinculados con la aplicación de esta prueba son indicativos de su mayor efectividad para detectar datos realmente útiles.

Con el fin de tener una primera aproximación de los efectos de las pruebas en la forma de los modelos, a continuación se revisan un ejemplo en el que la prueba t por pares produce un menor error que la de los signos, uno en que los errores son iguales y el único en el que el error de la prueba t es mayor.

4.2 Familia 00-ECb-DCb-PMa

Para esta familia, de hecho, se presenta el caso de mayor disminución del error en la estimación de los modelos con la prueba t. En primera instancia, y con referencia en el resumen de resultados de las dos pruebas de la Tabla 4.2, se advierte una reducción del ciclo con datos útiles de 2004 - 2008 con la prueba de los signos a 2005 - 2008 con la prueba t por pares.

Tabla 4.2 Valores P de la familia 00-ECb-DCb-PMa.

Grupo	Prueba	Valor P por series de mediciones consecutivas								
		04 - 05	05 - 06	06 - 07	07 - 08	08 - 09	09 - 10	10 - 11	11 - 12	12 - 13
1993-2009-2012	Signos	0.063	0.081	1.000	1.000	0.000	0.985	0.063	0.738	0.001
1993-2010-2012 (2)		0.000	0.958	1.000	0.500	0.019	1.000	0.151	0.042	0.886
1993-2009-2012	t por pares	0.001	0.031	1.000	1.000	0.000	1.000	0.536	0.176	0.195
1993-2010-2012 (2)		0.000	0.987	1.000	0.503	0.000	1.000	0.506	0.000	0.978

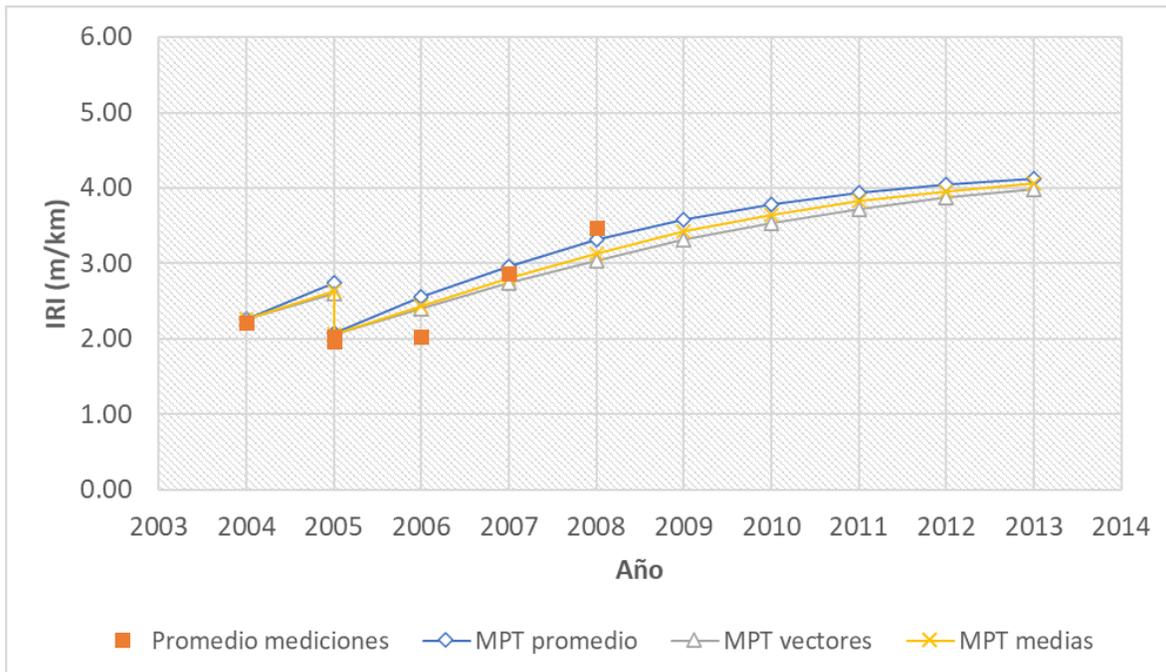


Figura 4.1 Valor esperado, familia 00-ECb-DCb-PMa, prueba de los signos.

Vinculado con lo anterior, la representación del modelo de la prueba de los signos de la Figura 4.1, muestra que, para el periodo de los datos rechazados por la prueba t (datos de 2004 - 2006 del segundo grupo), la evidencia gráfica del proceso de deterioro es bastante endeble.

En contraste, la Figura 4.2 revela que si no se toman en cuenta los datos de esos años, el modelo se ajusta mejor a las mediciones efectuadas, lo que, a su vez, se traduce en la disminución del error observada.

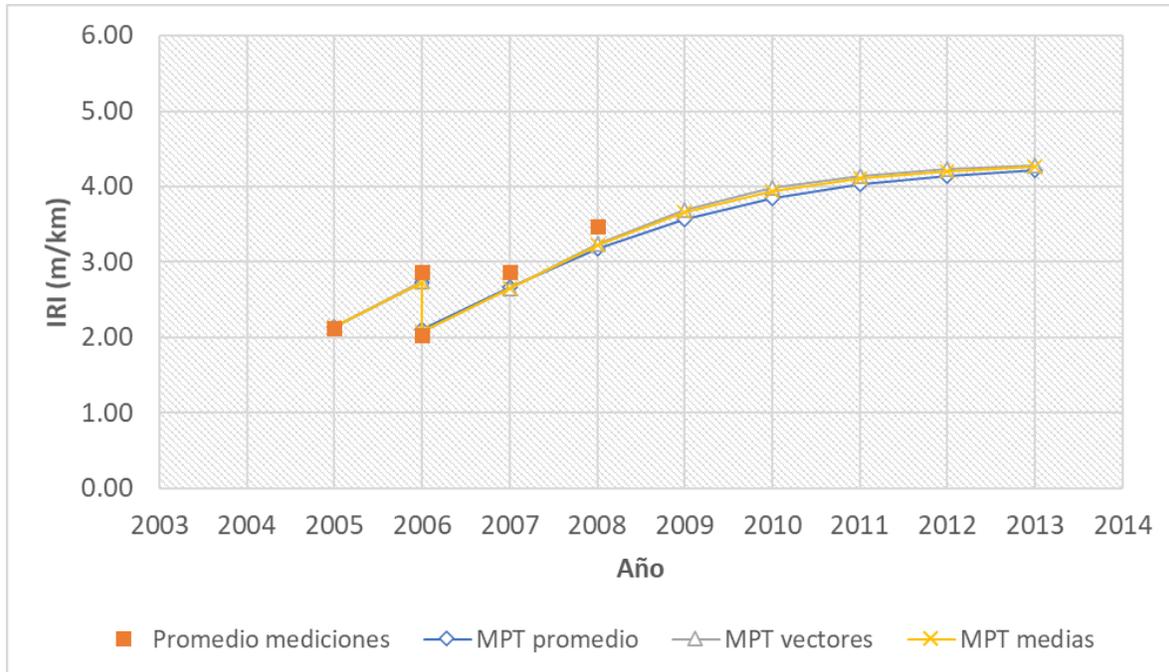


Figura 4.2 Valor esperado, familia 00-ECb-DCb-PMa, prueba t por pares.

4.3 Familia 00-ECm-DCa-PMa

En este caso, la duración de los ciclos utilizados para la modelización es la misma, a pesar de que, como lo presenta la Tabla 4.3, la prueba t rechaza los datos del primer grupo correspondientes al periodo 2005 - 2006.

Tabla 4.3 Valores P de la familia 00-ECm-DCa-PMa.

Grupo	Prueba	Valor P por series de mediciones consecutivas								
		04 - 05	05 - 06	06 - 07	07 - 08	08 - 09	09 - 10	10 - 11	11 - 12	12 - 13
2002-2007-2010	Signos	0.903	0.382	1.000	0.000	0.460	0.002	1.000		
2002-2007-2007		0.556	0.336	1.000	0.955	0.000	0.389	1.000		
2002-2007-2010	t por pares	0.683	0.023	1.000	0.000	0.280	0.116	1.000		
2002-2007-2007		0.534	0.110	1.000	0.947	0.000	0.189	1.000		

De acuerdo con las Figuras 4.3 y 4.4 este hecho no influye en la forma de los modelos, como tampoco tiene ningún efecto en el error, que es el mismo para las dos pruebas.

Nuevamente, se observa una evidencia gráfica que no es compatible con el deterioro del pavimento (periodo 2004 - 2006), aunque esto no queda expresado en términos de significación estadística ni aun con la prueba t por pares.

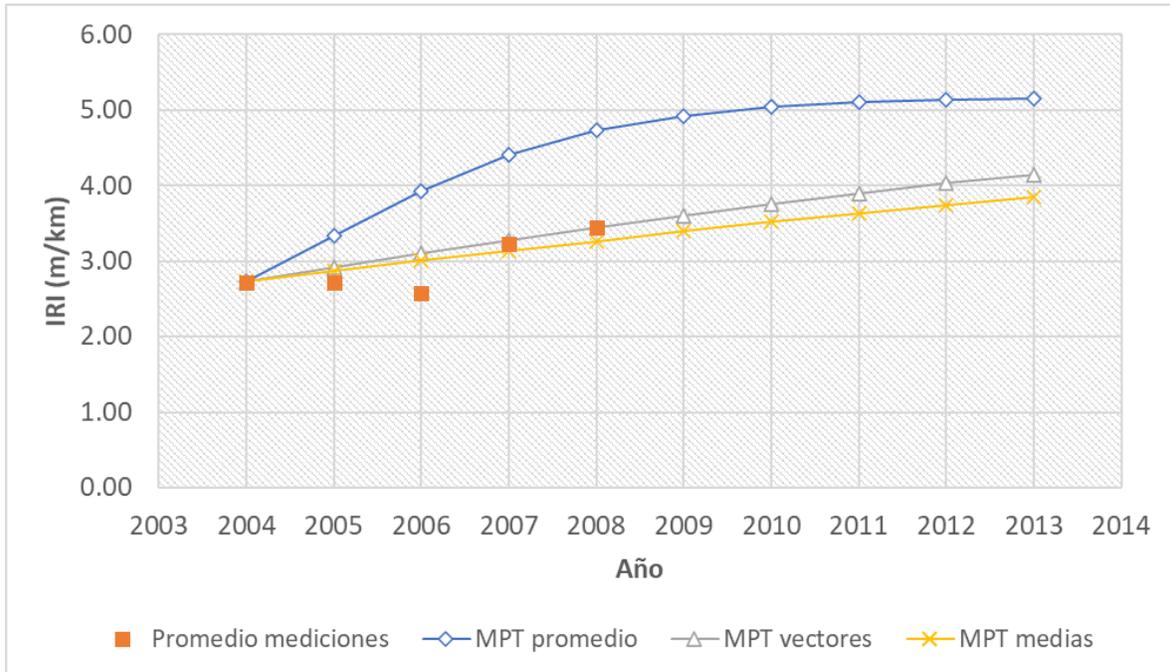


Figura 4.3 Valor esperado, familia 00-ECm-DCa-PMa, prueba de los signos.

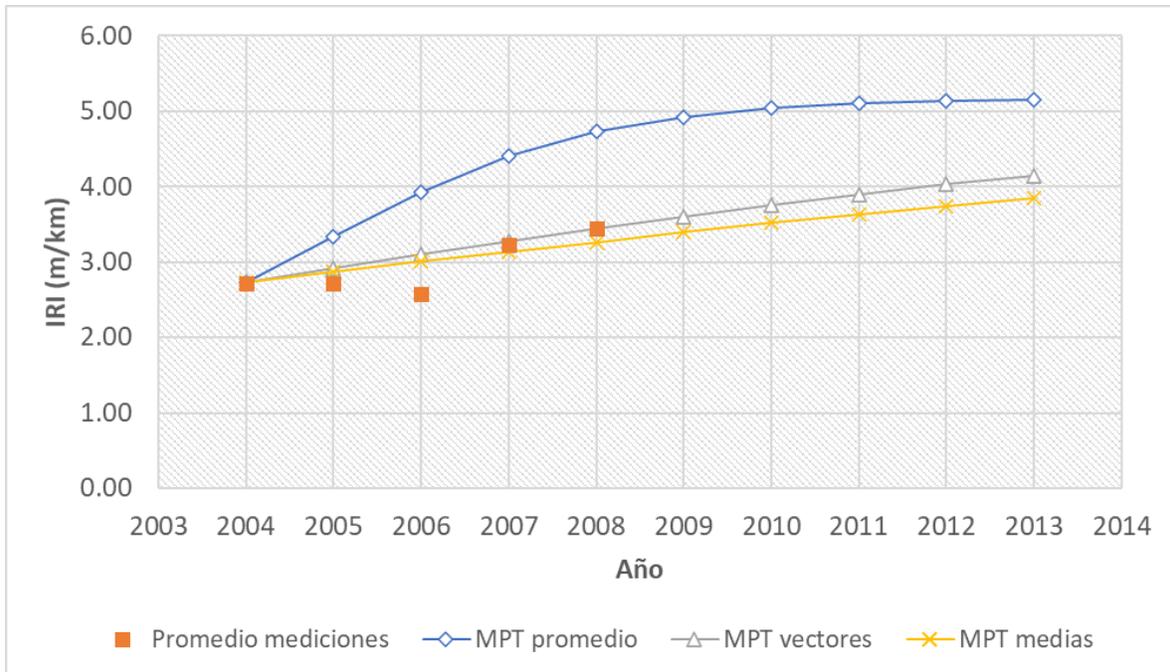


Figura 4.4 Valor esperado, familia 00-ECm-DCa-PMa, prueba t por pares.

4.4 Familia 00-ECm-DCm-PMa

A esta familia corresponde el único ejemplo en el que el error generado con modelos derivados de la aplicación de la prueba t es mayor al que se obtiene con la prueba de los signos.

Lo primero que llama la atención aquí, es que la familia abarca un conjunto de grupos muy numeroso. Por sí mismo, este hecho podría ser una fuente importante de error, ya que el riesgo de mayor heterogeneidad de los datos crece con la diversidad de los orígenes. En general, la prueba t por pares rechaza más series anuales consecutivas que la prueba de los signos, sin embargo, este comportamiento algo distinto no parece tener ningún efecto en la forma de los modelos, de acuerdo con las Figuras 4.5 y 4.6, ni tampoco en el nivel de error que, en apariencia, se encuentra mayormente condicionado por la posible variabilidad de las observaciones individuales.

Tabla 4.4 Valores P de la familia 00-ECm-DCm-PMa.

Grupo	Prueba	Valor P por series de mediciones consecutivas								
		04 - 05	05 - 06	06 - 07	07 - 08	08 - 09	09 - 10	10 - 11	11 - 12	12 - 13
2009-2009-2011	Signos	0.127	0.000	1.000	0.352	0.000	0.002	0.000	1.000	1.000
1993-1993-1993		0.271	0.000	1.000	1.000	1.000	0.935	0.763	0.000	1.000
1993-1993-2003		0.698	0.000	1.000	1.000	1.000	0.477	0.657	0.000	1.000
1993-2007-2007		0.057	0.092	1.000	1.000	1.000	0.990	1.000	0.000	1.000
2002-2012-2012		0.747	0.000	0.969	0.395	0.031	0.000	1.000		
2002-2002-2002		0.003	0.285	0.329	0.999	0.107	0.077	0.979		
2002-2011-2011		0.047	0.173	0.916	1.000	0.000	0.138	0.047		
2002-2002-2010		0.982	0.998	0.186	0.000	1.000	1.000	0.000		
2009-2009-2009 (1)		0.997	0.008	0.984	0.240	0.941	0.000	1.000		
2009-2009-2009 (2)		0.007	0.000	1.000	0.473	0.000	0.000	1.000		
2009-2009-2009 (3)		0.343	0.002	0.523	0.737	0.999	0.000	1.000		
2009-2009-2009 (4)		0.055	0.000	1.000	0.748	0.000	0.000	1.000		
2010-2010-2010		0.761	0.055	0.945	0.857	0.999				
2012-2012-2012		0.041	0.022	1.000	0.768	0.973	0.739	0.998		
2009-2009-2011		t por pares	0.215	0.000	1.000	0.545	0.000	0.016	0.000	1.000
1993-1993-1993	0.691		0.000	1.000	1.000	1.000	1.000	1.000	0.000	1.000
1993-1993-2003	0.223		0.000	1.000	1.000	1.000	0.968	0.994	0.000	1.000
1993-2007-2007	0.012		0.002	1.000	1.000	1.000	1.000	1.000	0.000	1.000
2002-2012-2012	0.770		0.000	0.996	0.625	0.002	0.001	1.000		
2002-2002-2002	0.001		0.296	0.795	0.934	0.000	0.141	0.975		
2002-2011-2011	0.037		0.671	0.991	1.000	0.000	0.108	0.048		
2002-2002-2010	0.998		1.000	0.037	0.000	1.000	1.000	0.000		
2009-2009-2009 (1)	0.991		0.000	0.996	0.005	0.647	0.000	0.999		
2009-2009-2009 (2)	0.000		0.000	1.000	0.666	0.000	0.000	1.000		
2009-2009-2009 (3)	0.626		0.000	0.958	0.233	0.855	0.000	1.000		
2009-2009-2009 (4)	0.060		0.001	1.000	0.933	0.000	0.000	1.000		
2010-2010-2010	0.543		0.141	0.999	0.647	0.621				
2012-2012-2012	0.018		0.015	1.000	0.668	0.025	1.000	0.977		

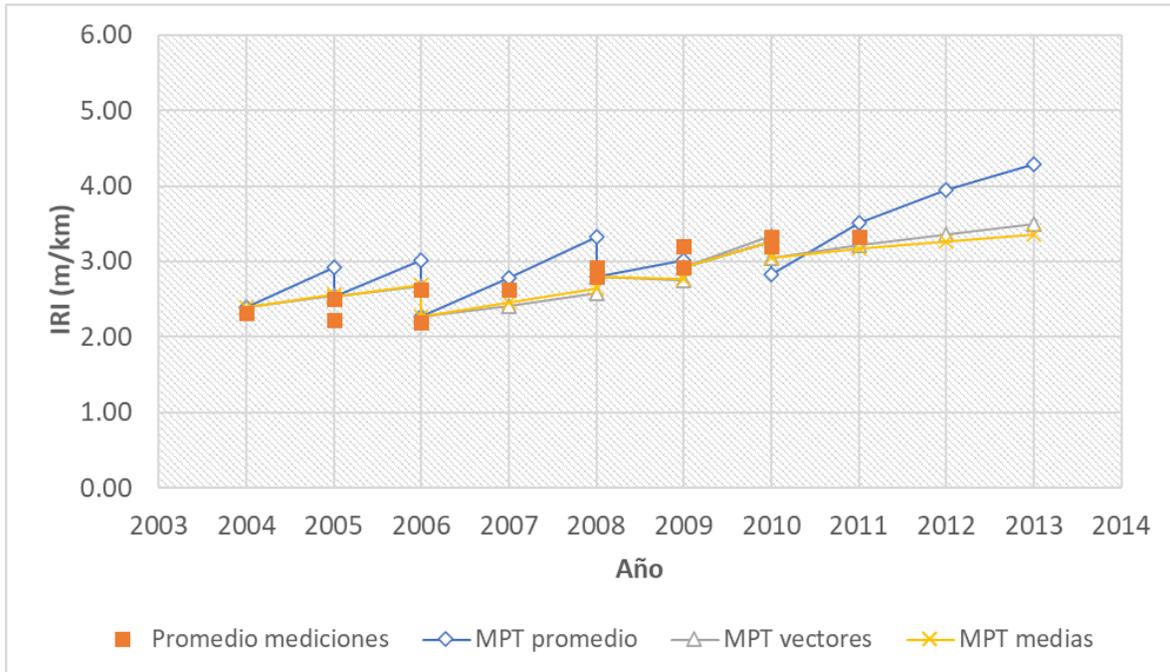


Figura 4.5 Valor esperado, familia 00-ECm-DCm-PMa, prueba de los signos.

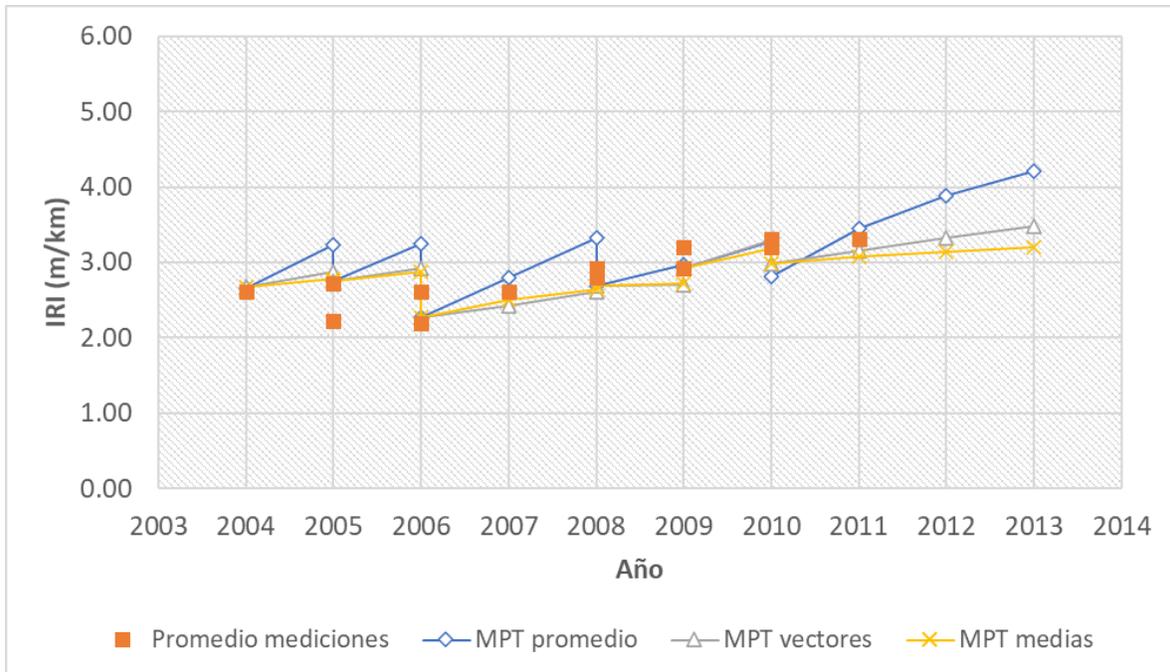


Figura 4.6 Valor esperado, familia 00-ECm-DCm-PMa, prueba t por pares.

5. Conclusiones

A partir de los resultados obtenidos en el trabajo descrito en este documento, pueden formularse las siguientes conclusiones:

- Se ha producido un software operativo con respecto a los dos principales procesos contemplados en la metodología de modelización desarrollada en el IMT: aplicación de pruebas de inferencia estadística para delimitar ciclos de deterioro y generación de modelos basados en cadenas de Márkov y en información histórica.
- Aunque, mediante el software desarrollado, se aplican algunos procesos complejos, el sistema de cómputo como tal es relativamente sencillo, ya que utiliza una sola entidad de datos en modo lectura y el número de procesos que implementa es bajo. Como consecuencia, no se aplicaron técnicas formales de análisis y diseño de programas.
- Se ha utilizado *Visual Basic for Applications (VBA)* de Excel como plataforma de desarrollo porque tanto los datos de entrada como los resultados del software tienen mayormente un formato tabular, además de que esta plataforma provee una importante biblioteca de funciones matemáticas, estadísticas y gráficas que pueden utilizarse durante la escritura del código.
- La prueba t por pares es aplicable con los datos disponibles, ya que, aunque estos no cumplen con el requisito de ajustarse a una distribución normal, sí tienen otras características reportadas como mínimas necesarias en la literatura técnica internacional para que la prueba pueda aplicarse. Tales características se resumen en que la distribución de los datos sea continua, unimodal y “razonablemente simétrica”.
- La comparación de los resultados producidos por las pruebas t por pares y de los signos permiten afirmar que el uso de la primera generalmente redundante en un menor error y en una mayor capacidad para reunir la evidencia estadística necesaria para rechazar la hipótesis nula cuando esta es, en efecto, falsa y, con ello, evitar la ocurrencia de errores tipo II.
- En esta versión del software no se contempló la codificación de los procesos de clasificación de tramos en familias y grupos ni de preparación de las series históricas. La razón es que, por el momento, las mediciones de la red federal de carreteras se encuentran almacenadas en una colección de decenas de miles de archivos planos, lo que dificulta considerablemente el acceso a la información y su procesamiento.

- El desarrollo de modelos de deterioro de pavimentos en México representa una vasta área de oportunidad, por lo que conviene apoyar cualesquier iniciativas de estudios de campo, análisis de información histórica y desarrollos teóricos relacionados con este tema.

Bibliografía

Arriaga, M. C., Garnica, P. & Rico, A., 1998. *Índice Internaciona de Rugosidad de la red carretera de México*, San Fandila, Qro.: Instituto Mexicano del Transporte.

ASTM, 2012. *ASTM E867 - 12: Standard Terminology Relating to Vehicle-Pavement Systems*. West Conshohocken(PA): ASTM International.

Brown, S., 2020-1. *Measures of Shape: Skewness and Kurtosis*. [En línea]
Available at: <https://brownmath.com/stat/shape.htm>
[Último acceso: 8 marzo 2021].

Brown, S., 2020-2. *Ch 10. Hypothesis Tests / SWT*. [En línea]
Available at: <https://brownmath.com/swt/chap10.htm>
[Último acceso: 10 Marzo 2021].

Bulmer, M. G., 1979. *Principles of Statistics*. Nueva York(NY): Dover Publications, Inc..

Buthmann, A., 2015. *Isix Sigma*. [En línea]
Available at: <http://www.isixsigma.com/tools-templates/normality/dealing-non-normal-data-strategies-and-tools/>
[Último acceso: 22 Diciembre 2020].

Butt, A. A., Shahin, M. Y., Carpenter, S. H. & Carnahan, J. A., 1994. *Application of Markov Process to a Pavement Management Systems at Network Level*, San Antonio, Texas, EUA: Transportation Research Board.

Corro, S., Magallanes, R. & Prado, G., 1981. *Instructivo para diseño estructural de pavimentos flexibles para carreteras. Series del Instituto de Ingeniería no. 444*, Ciudad de México: Instituto de Ingeniería, UNAM.

Fylstra, D., Lasdon, L., Watson, J. & Waren, A., 1998. *Design and Use of the Microsoft Excel Solver*, Nevada: Institute for Operations Research and the Management Sciences.

Garnica, P. & Hernández, R., 2013. *Manual de usuario IMT-PAVE 1.1: Manual de Usuario*, San Fandila, Qro.: Instituto Mexicano del Transporte.

Gillespie, T. D., Sayers, M. & Segel, L., 1980. *NCHRP Report 228: Calibration of Response-Type Road Roughness Measuring Systems*, Washington, DC, EUA: TRB, National Research Council.

Lytton, R. L., 1987. *Concepts of Pavement Performance Prediction and Modeling*. Toronto, Canada, s.n.

Martínez, J. J. y otros, 2017. *Manual estadístico del sector transporte 2017*, San Fandila, Qro.: Instituto Mexicano del Transporte.

Minitab, 2010. *Statistical Inference and t-Tests*. s.l.:Minitab Inc..

Minitab, 2019-1. *Data considerations for Paired t*. [En línea]
Available at: <https://support.minitab.com/en-us/minitab-express/1/help-and-how-to/basic-statistics/inference/how-to/two-samples/paired-t/before-you-start/data-considerations/>
[Último acceso: 7 marzo 2021].

Minitab, 2019-2. *Methods and formulas for 1-Sample Sign*. [En línea]
Available at: <https://support.minitab.com/en-us/minitab-express/1/help-and-how-to/basic-statistics/inference/how-to/one-sample/1-sample-sign/methods-and-formulas/methods-and-formulas/#p-value-for-the-exact-method>
[Último acceso: 7 marzo 2021].

Minitab, 2019-3. *Methods and formulas for Paired t*. [En línea]
Available at: <https://support.minitab.com/en-us/minitab-express/1/help-and-how-to/basic-statistics/inference/how-to/two-samples/paired-t/methods-and-formulas/methods-and-formulas/#confidence-interval-ci>
[Último acceso: 8 marzo 2021].

Montgomery, D. C. & Runger, G. C., 2003. *Applied Statistics and Probability for Engineers*. Tercera ed. Nueva York(NY): John Wiley & Sons, Inc.

Ningyuan, L., Wei-Chau, X. & Haas, R., 1996. Reliability-Based Processing of Markov Chains for Modeling Pavement Network Deterioration. *Transportation Research Record*, Issue 1524, pp. 203-213.

NIST/SEMATECH, 2013. *7.1.3.1. Critical values and p values*. [En línea]
Available at: <https://www.itl.nist.gov/div898/handbook/prc/section1/prc131.htm>
[Último acceso: 11 marzo 2021].

NIST, 2015. *Sign Test*. [Online]
Available at:
<https://www.itl.nist.gov/div898/software/dataplot/refman1/auxillar/signtest.htm>
[Accessed 6 03 2021].

NIST, 2018. *T Test*. [En línea]
Available at:
https://www.itl.nist.gov/div898/software/dataplot/refman1/auxillar/t_test.htm
[Último acceso: 7 marzo 2021].

Ortiz-García, J. J., Costello, S. B. & Snaith, M. S., 2006. Derivation of Transition Probability Matrices for Pavement Deterioration Modeling. *ASCE Journal of Transportation Engineering*, Febrero.

Robinson, R., Danielson, U. & Snaith, M., 1998. *Road Maintenance Management*. Basingstoke(Hampshire): Palgrave.

Romero, J. A. & Lozano, A., 1995. *La respuesta dinámica de un cuarto de carro y el Índice Internacional de Rugosidad. Publicación Técnica no. 67.*, San Fandila, Qro.: Instituto Mexicano del Transporte.

Sayers, M. W., Gillespie, T. D. & Paterson, W. D. O., 1986-1. *Guidelines for conducting and calibrating road roughness measurements*, UK: Banco Mundial.

Sayers, M. W., Gillespie, T. D. & Queiroz, C. A. V., 1986-2. *The International Road Roughness Experiment*, Washington, D.C., EUA: The World Bank.

SCT, 2018. *Programa de Auscultación de la Red Carretera Federal: Memoria documental*, Ciudad de México: Secretaría de Comunicaciones y Transportes.

Solorio, R., 2018. *Aplicación de Cadenas de Márkov para modelizar la evolución del IRI en la red federal de carretera*, Chihuahua, Chih.: UACH.

The European Mathematical Society, 2020. *Student test*. [En línea] Available at: https://encyclopediaofmath.org/wiki/Student_test [Último acceso: 7 marzo 2021].



COMUNICACIONES

SECRETARÍA DE COMUNICACIONES Y TRANSPORTES



Km 12+000 Carretera Estatal 431 "El Colorado Galindo"
Parque Tecnológico San Fandila, Mpio. Pedro Escobedo,
Querétaro, México. C.P. 76703
Tel: +52 (442) 216 97 77 ext. 2610
Fax: +52 (442) 216 9671

publicaciones@imt.mx

<http://www.imt.mx/>