



## **SERIES TEMPORALES CON EXCEL**

**Máster de Dirección de Empresas Turísticas**

**Métodos de Investigación Aplicados al Turismo**

- 1. INTRODUCCIÓN**
- 2. REPRESENTACIÓN GRÁFICA**
- 3. DETERMINACIÓN DEL ESQUEMA DE INTEGRACIÓN DE LAS COMPONENTES**
- 4. DESCOMPOSICIÓN ESTACIONAL**
- 5. PREDICCIONES A CORTO PLAZO**
- 6. REFERENCIAS**

# 1. INTRODUCCIÓN

En el capítulo 8 se aborda el análisis y representación gráfica de las series temporales siguiendo el **enfoque clásico de descomposición**. Este método supone que la serie temporal está formada por cuatro componentes (tendencia, ciclo, estacionalidad y componente irregular), de tal forma que el método funciona extrayendo en primer lugar cada una de dichas componentes para analizarlas por separado –lo que permite adquirir un conocimiento sobre las series- y después, volver a combinarlas para formar las predicciones. En esta práctica, aplicaremos dicha técnica utilizando el software de aplicación horizontal **Excel**. Nos centraremos en el cálculo de la componente estacional (**Índices de Variación Estacional**), la obtención de la **serie desestacionalizada** y realizaremos una **predicción a corto plazo** para la variable estudiada.

El fichero de datos utilizado para ilustrar las distintas aplicaciones de este documento se denomina **MIAT\_MDET.xls** y contiene la serie trimestral del número de viajeros alojados en establecimientos hoteleros de la provincia de Huelva, extraída de *la Encuesta de Ocupación de Establecimientos Turísticos*, que tiene periodicidad mensual y ha sido totalizada por trimestres. El periodo abarca desde el primer trimestre del año 2000 al segundo trimestre de 2013. *Dichos datos pueden ser descargados de la base de datos INEbase del INE.*

Para llevar a cabo el análisis clásico descriptivo de una serie temporal habrá que seguir los siguientes pasos:

- Introducir datos
- Representación gráfica de la serie
- Determinación del modelo adecuado (aditivo o multiplicativo)
- Determinación de la estacionalidad

The screenshot shows the INEbase website interface. The browser address bar displays the URL: `www.ine.es/faxi/tabla.do?path=/t38/bme2/t04/a082/10/&file=1202003.px`. The page title is "Principales resultados sobre energía" and the sub-section is "Industria".

Under the heading "Consumo de energía final" (Unidades: Miles de Tep), there is a section "Seleccione valores a consultar:" with two dropdown menus:

- Energía:** A list of energy sources including TOTAL, Carbón, Gases derivados del carbón, Productos petrolíferos, Gas natural y manufacturado, Electricidad, and Energías renovables. "Energías renovables" is selected.
- Periodo:** A list of time periods from 2011M08 to 2012M03. "2011M08" is selected.

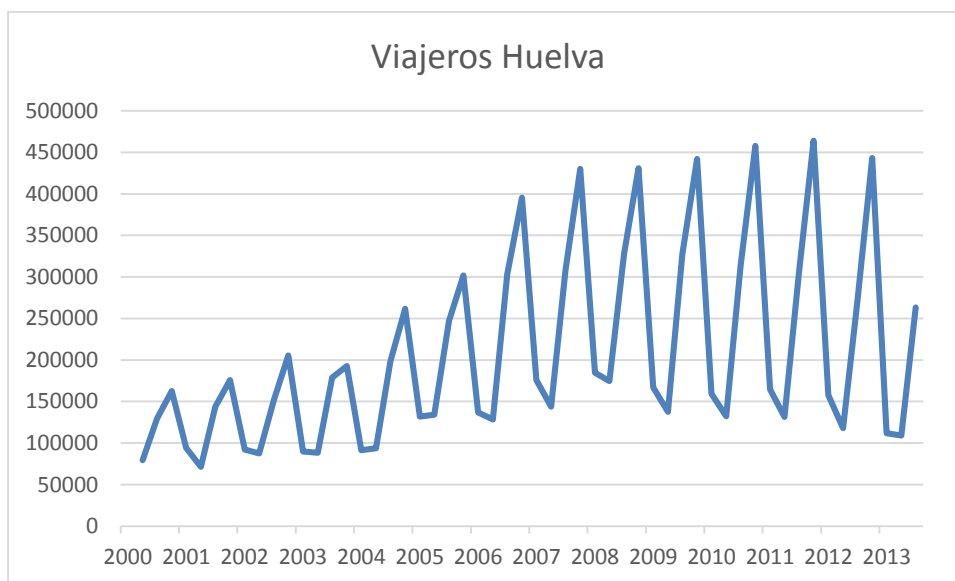
Below these menus, there is a section "Elija cómo quiere visualizar sus datos:" with a "Variables en filas" dropdown menu showing "Periodo" selected.

- Obtención de la serie desestacionalizada
- Predicción a corto plazo

## 2. REPRESENTACIÓN GRÁFICA DE LA SERIE

El punto de partida en el estudio de una serie temporal es su representación gráfica, lo cual persigue identificar patrones de comportamiento regulares que permitan plantear hipótesis sobre la estacionalidad y el ciclo, el esquema según el cual se combinan las componentes de la serie temporal o la existencia de *outliers* o valores atípicos en la serie temporal que adviertan de errores en la recolección de los datos o de fenómenos imprevistos de un impacto excepcionalmente grande en los valores de la variable.

Para llevar a cabo la representación gráfica de una serie temporal, en Excel, seleccionamos **Insertar/Gráfico/Líneas**. El resultado se muestra a continuación:



Fuente: INE

*En el gráfico se aprecia una marcada componente estacional y una variabilidad que aumenta al aumentar la tendencia de la serie. Sistemáticamente, los meses de verano muestran un mayor número de viajeros que los meses de invierno. La tendencia no se mantiene creciente durante el período analizado, sino que hay una fase expansiva, con una estabilización y, finalmente, un decrecimiento.*

### 3. DETERMINACIÓN DEL ESQUEMA DE INTEGRACIÓN DE LAS COMPONENTES

Una vez analizada gráficamente la serie, el paso siguiente consiste en determinar cuál de los esquemas de integración es el que mejor se ajusta a la serie en estudio. Las posibilidades que se consideran habitualmente son las siguientes:

- Aditivo:  $Y = T + C + e + I$
- Multiplicativo:  $Y = T * C * e * I$
- Mixto:  $Y = T * C * e + I$

Nos centraremos en los dos primeros. Los métodos para determinar el esquema de integración son los siguientes:

- Métodos gráficos:
  1. Gráfico de secuencia
  2. Gráfico Media/desviaciones típicas
- Métodos numéricos:
  3. Variabilidad de las diferencias y cocientes estacionales

1. **Análisis del gráfico de secuencia.** Se basa en el análisis visual de la serie. Es útil en casos extremos, y a menudo se aplican los otros métodos para decidir el esquema de integración. Básicamente, si las oscilaciones alrededor de la tendencia se mantienen constantes, correspondería al modelo aditivo; mientras que si éstas se van amplificando (o disminuyendo) a medida que crece (o disminuye) la tendencia, se ajustaría a un modelo multiplicativo.

*En nuestro ejemplo, tal como se comentó en el apartado 2, las oscilaciones parece que se van amplificando a medida que aumenta la tendencia de la serie, por lo que sería más adecuado considerar un modelo multiplicativo.*

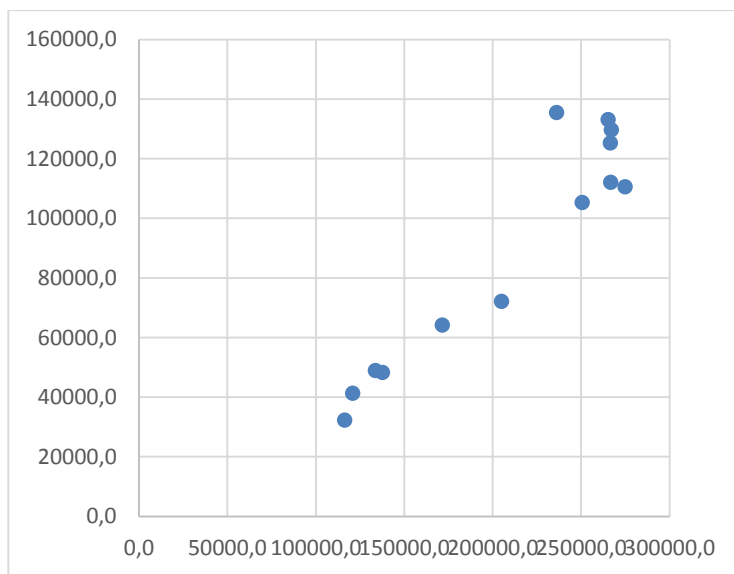
2. **Gráfico media/desviación típica.** En este caso, mediante la opción de tablas dinámicas, creamos una tabla de doble entrada en la que recogemos las observaciones de la serie clasificadas por filas según la periodicidad de la serie (mes, trimestre, ...) y por columnas, los años. De este modo, podremos calcular para cada año, la media y desviación típica de los valores de la variable en cada estación.

Insertar/Tabla dinámica



	A	B	C	D	E	F	G
1							
2							
3	Promedio de Viajeros Huelva	Etiquetas de columna					
4	Etiquetas de fila		2000	2001	2002	2003	2004
5	1		79488	71328	87457	88119	93868
6	2		129404	143607	151796	179014	198395
7	3		162578	175947	205618	192598	261779
8	4		94223	92307	89840	91402	131816
9	(en blanco)						
10	Total general		116423,25	120797,25	133677,75	137783,25	171464,5
11							
12	Media		116423,3	120797,3	133677,8	137783,3	171464,5
13	Desviación		32232,8	41293,7	48892,2	48276,3	64177,8
14							

A continuación representamos las medidas obtenidas (media y desviación típica) en un gráfico de dispersión; de modo que, si la nube de puntos es creciente, el modelo adecuado sería el multiplicativo; y en caso contrario, el modelo aditivo. *En nuestro ejemplo, dicho gráfico confirma la elección del modelo multiplicativo como el más idóneo.*



## 4. DESCOMPOSICIÓN ESTACIONAL.

El procedimiento para determinar los coeficientes de estacionalidad (*e*) o Índices Generales de Variación Estacional Ajustados (**IGVEA**) difiere según sea el esquema de integración de los componentes. En lo que sigue, desarrollaremos los cálculos para el caso del modelo multiplicativo con datos mensuales, con objeto de seguir con nuestro ejemplo.

Los pasos son:

1. Calculamos una serie de medias móviles para cada  $p=4$  observaciones, al tratarse de datos trimestrales. El primer valor de la media móvil en nuestra serie es **=promedio(c3:c6)** y se situaría en la celda **D5**. Una vez calculada, arrastraríamos la fórmula hasta la celda **D55**.

	A	B	C	D	E	F
1	AÑO	TRIM	Viajeros Huelva	MMD	MMC	
2						
3	2000	1	79488			
4	2000	2	129404			
5	2000	3	162578	=PROMEDIO(C3:C6)		
6	2000	4	94223	114383		
7	2001	1	71328	117934		
8	2001	2	143607	121276		
9	2001	3	175947	120797		
10	2001	4	82227	121820		

Como  $p$  es par, para centrar la serie, calculamos nuevamente medias móviles, pero, en este caso, de orden 2. Así, el primer valor de esta serie, ya centrada, se sitúa en la celda **E5** y su expresión numérica sería **=promedio(D5:D6)**.

	A	B	C	D	E	F
1	AÑO	TRIM	Viajeros Huelva	MMD(4)	MMC	
2						
3	2000	1	79488			
4	2000	2	129404			
5	2000	3	162578	116423	=PROMEDIO(D5:D6)	
6	2000	4	94223	114383	116159	
7	2001	1	71328	117934	119605	
8	2001	2	143607	121276	121037	

Arrastraríamos la fórmula hasta la celda **E54**. Y así, obtendríamos las **medias móviles centradas**, que recogen la componente extraestacional (tendencia-ciclo):

$$mmc = T * C$$

- Obtención de los **Índices Específicos de Variación Estacional -IEVE**. Al tratarse de un modelo multiplicativo, se obtienen dividiendo la serie original entre la tendencia-ciclo (mmc). La nueva serie recoge la estacionalidad y la componente irregular o residuo.

$$IEVE = \frac{Y}{mmc} = \frac{T * C * e * I}{T * C} = e * I$$

El primer valor de esta nueva serie se sitúa en la celda **F3** y para obtener su valor escribimos **=C5/E5\*100** y arrastramos hasta **F54**.

	A	B	C	D	E	F	G
1	AÑO	TRIM	Viajeros Huelva	MMD(4)	MMC	IEVE	
2							
3	2000	1	79488				
4	2000	2	129404				
5	2000	3	162578	116423,25	115403,25	=C5/E5*100	
6	2000	4	94223	114383,25	116158,63	81,12	
7	2001	1	71328	117934,00	119605,13	59,64	
8	2001	2	143607	121276,25	121036,75	118,65	

- Obtención de los **Índices Generales de Variación Estacional -IGVE**. Se calculan promediando los IEVE para cada estación (trimestre). De esta forma, se elimina el efecto de la componente irregular. Para ello podemos insertar en la celda **G5** la fórmula del promedio condicional dada por la expresión **=PROMEDIO.SI(B5:B54;B5;F5:F54)**. Antes de arrastrar hasta la celda **G54**, acotamos las referencias de las áreas con el signo \$ para fijarlas al arrastrar: **=PROMEDIO(\$B\$5:\$B\$54;B5;\$F\$5:\$F\$54)**, de forma que sólo cambiará el segundo argumento, la celda en la que figura el trimestre que corresponde.

	A	B	C	D	E	F	G
1	AÑO	TRIM	Viajeros Huelva	MMD(4)	MMC	IEVE	IGVE
2							
3	2000	1	79488				
4	2000	2	129404				
5	2000	3	162578	116423,25	115403,25	140,88	=PROMEDIO.SI(\$B\$5:\$B\$54;B5;\$F\$5:\$F\$54)
6	2000	4	94223	114383,25	116158,63	81,12	65,36
7	2001	1	71328	117934,00	119605,13	59,64	57,57
8	2001	2	143607	121276,25	121036,75	118,65	118,99
9	2001	3	175047	120707,25	122812,25	142,26	157,71

Promedio de	Año					IGVE	
Mes		2008	2009	2010	2011	2012	
1			117,2713	118,8	118,6714393		118,25
2			113,5823	112,1	108,2080925		111,31
3			111,1724	113,6	110,6848421		111,83
4			108,9304	95,93	94,39948203		99,75
5			99,58609	97,73	95,55806816		97,62
6			92,22368	89,58	92,28527483		91,36
7			94,82776	88,38	83,32793784		88,85
8			77,52577	78,72	90,44153324		82,23
9			88,94926	87,9	92,8669521		89,91
10		95,7	95,74099	91,57			94,34
11		90,7886	104,9986	114,9			103,56
12		88,0968	107,1388	121,3			105,52
						promedio	99,54

A continuación se calcula la media aritmética de todos los IGVE.

$$\overline{IGVE} = \frac{IGVE_{1T} + IGVE_{2T} + \dots + IGVE_{4T}}{4}$$

- Finalmente, hallamos los **Índices Generales de Variación Estacional Ajustados -IGVEA**, cuya expresión matemática sería el cociente entre IGVE de cada mes y el promedio de los IGVE (multiplicándolo por 100 para que queden expresados en tanto por ciento).

$$IGVEA_{1T} = \frac{IGVE_{1T}}{\overline{IGVE}} * 100$$

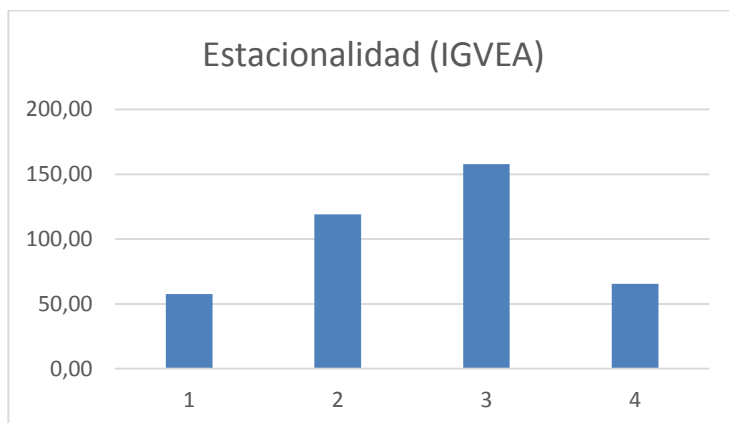
En Excel, para obtener el IGVEA del tercer trimestre de 2000, celda **H5**, se ha aplicado la fórmula **=G5/(PROMEDIO(\$G\$5:\$G\$8))\*100**. Para el resto de los trimestres, basta con arrastrar dicha fórmula, completando los dos valores iniciales y finales, dado que son idénticos para cada trimestre de año en año.

AÑO	TRIM	Viajeros Huelva	MMD(4)	MMC	IEVE	IGVE	IGVEA
2000	1	79488					57,63
2000	2	129404					119,10
2000	3	162578	116423,25	115403,25	140,88	157,71	=G5/(PROMEDIO(\$G\$5:\$G\$8))*100
2000	4	94223	114383,25	116158,63	81,12	65,36	65,42
2001	1	71328	117934,00	119605,13	59,64	57,57	57,63
2001	2	143607	121276,25	121036,75	118,65	118,99	119,10

Los resultados confirman el análisis gráfico del apartado 2: el mayor número de viajeros se manifiesta en el tercer trimestre – verano - de cada año, con un 157,86% sobre la media trimestral, mientras que el segundo – primavera – supone el 118,99% sobre a media. Tanto el último trimestre de cada año (65,36%) como en el primero (57,57%) los viajeros alojados en establecimientos

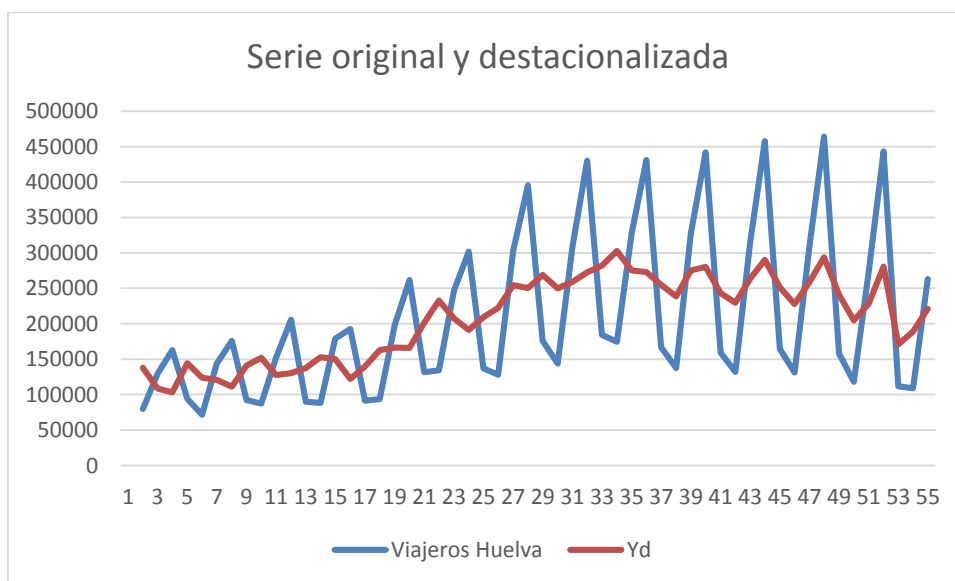


hoteleros no alcanzan le media trimestral, siendo en éste último en el que se dan las cifras más bajas.



- Finalmente, para la obtención de la **serie desestacionalizada (Yd)** eliminamos las variaciones estacionales dividiendo cada valor de la serie por el correspondiente IGVEA y multiplicando por 100, es decir, utilizando la expresión dada por  $=C3/H3*100$  en la celda I3 y arrastrando hasta I56.

	A	B	C	D	E	F	G	H	I
	AÑO	TRIM	Viajeros Huelva	MMD(4)	MMC	IEVE	IGVE	IGVEA	Yd
3	2000	1	79488					57,63	=C3/H3*100
4	2000	2	129404					119,10	108650,57
5	2000	3	162578	116423,25	115403,25	140,88	157,71	157,86	102991,93

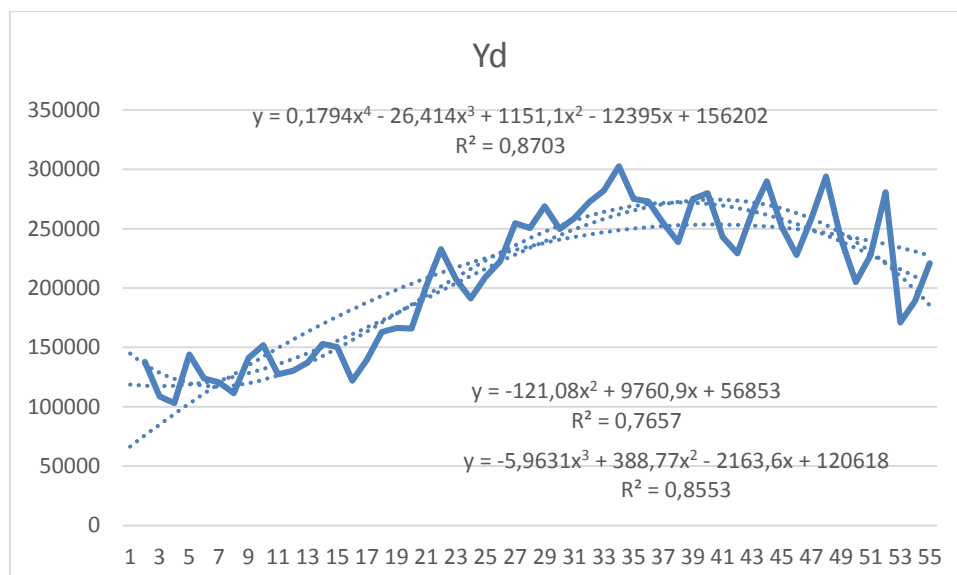


## 5. PREDICCIÓN A CORTO PLAZO

La serie desestacionalizada puede emplearse para efectuar predicciones. Para propósitos de predicción, el proceso se invierte. En lugar de separar la serie en sus componentes para examinarla, las componentes se vuelven a combinar para formar la predicción. Para ello, se proyectan para cada periodo futuro del que se quiere la predicción cada una de las componentes aisladas, y se usa el modelo multiplicativo o aditivo para formar el pronóstico. En general, para efectuar la predicción consideraremos la tendencia calculada y los valores estacionales.

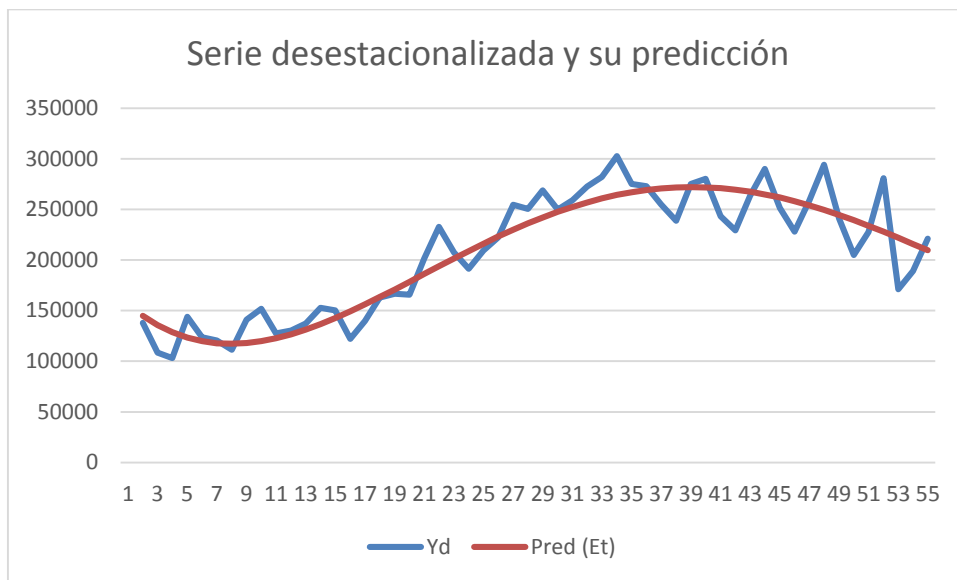
Para ello, seguiremos los siguientes pasos:

1º. A partir de la serie desestacionalizada, estimaremos la **tendencia** por el método de regresión lineal. Para ello, representamos gráficamente la serie desestacionalizada y, seguidamente, situándonos sobre la curva y pulsando el botón derecho del ratón, aparece un menú desplegable en el que seleccionamos la opción Agregar línea de tendencia. Existen muchas posibilidades, podemos probar con una recta y con modelos polinómicos de grados 2, 3, 4,.. En nuestro caso parece que el modelo que mejor representa el comportamiento de la serie es el polinómico de cuarto grado.



Definimos en la columna **J**, desde **J3** a **J58**, es decir, hasta el cuarto trimestre de 2013, una variable  $t=1, 2, 3, \dots$  que representa el período temporal, y que se emplea como variable independiente en el modelo de regresión, y considerando la ecuación del

polinomio de cuarto grado, podemos darle valores en la celda K3 mediante la expresión  $= (0,1794 * J3^4) - (26,414 * J3^3) + (1151,1 * J3^2) - (12395 * J3) + 156202$  que arrastraremos hasta K58, obteniendo las predicciones del componente extraestacional – tendencia y ciclo – para la serie analizada, que representamos junto a la serie desestacionalizada:

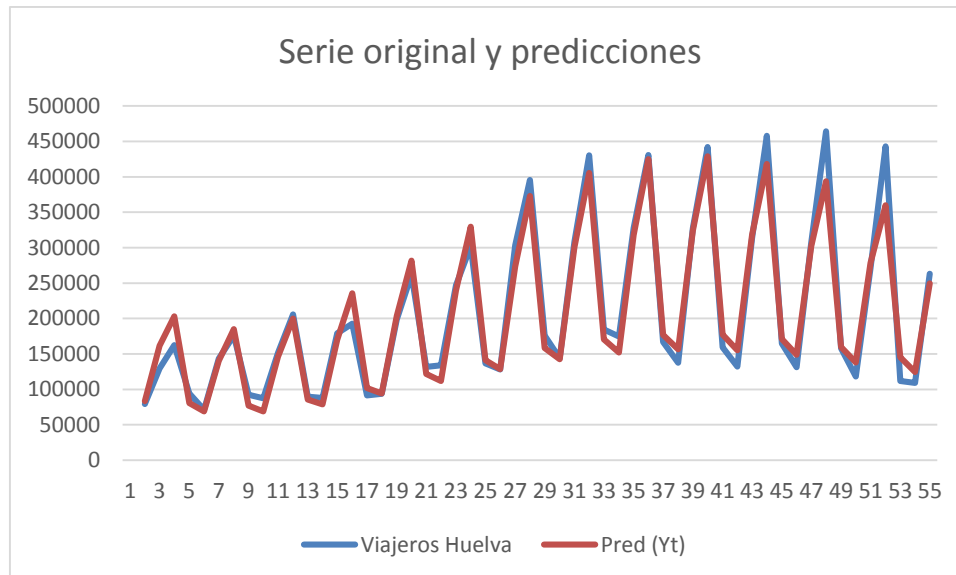


Y también frente a la serie original, evidenciándose que estas predicciones no incorporan aún el componente estacional



2. Para obtener predicciones de los valores de la serie, que incorporen los efectos estacionales basta con corregir los valores de la tendencia estimados con la estacionalidad previamente aislada. En nuestro ejemplo, multiplicaríamos el valor de la tendencia por el índice general de variación estacional ajustado del mes

correspondiente, contenido en la columna **H**, y dividiríamos por 100; es decir, para el primer trimestre de 2000, nos situamos en la celda **L3** e introducimos la fórmula dada por la expresión  $=K3*H3/100$ , arrastrando posteriormente hasta la celda **L56**.



La serie así obtenida incorpora las predicciones de la tendencia y el ciclo y los efectos estacionales, pudiendo comprobarse la relativa bondad de las predicciones “ex - ante” mediante la representación gráfica de esta nueva serie y la serie original. Hay ciertos desajustes en los últimos años que pueden deberse a un cambio en la estacionalidad como resultado de la crisis económica y la reducción de la temporada de apertura de bastantes hoteles para reducir costes.

Deliberadamente no hemos incluido en la serie original los valores correspondientes a los dos últimos trimestres de 2013, que fueron, respectivamente, 447.130 viajeros para el tercer trimestre y 138.607 viajeros alojados para el cuarto trimestre. Nuestras predicciones para dichos trimestres han sido, respectivamente de 321.309 viajeros y de 129.211 viajeros, por lo que hay una significativa subvaloración en el tercer trimestre y un ajuste bastante aproximado para el cuarto trimestre

## 6. REFERENCIAS

- J. E. Hanke y D.W. Wichern, *Business Forecasting*, 8ª ed. (Prentice Hall)
- E. Uriel, *Análisis de datos. Series temporales y análisis multivariante*. (Ed. AC)