

# RECAUDACIÓN DE INGRESOS BRUTOS EN LA PROVINCIA DE SANTA CRUZ, REPÚBLICA ARGENTINA: CARACTERIZACIÓN, ESTIMACIÓN Y PRONÓSTICO CON SERIES DE TIEMPO<sup>1</sup>

## COLLECTION OF GROSS INCOME IN THE PROVINCE OF SANTA CRUZ, ARGENTINE REPUBLIC: CHARACTERIZATION, ESTIMATION AND FORECAST WITH TIME SERIES

Juan Domingo, Arias-Montes<sup>2</sup>

Universidad Tecnológica Nacional Facultad Regional Santa Cruz

**Resumen:** La recaudación tributaria en los gobiernos subnacionales desempeña un papel importante en el proceso de formulación presupuestaria y de planificación estratégica. Aunque se conocen los factores determinantes, se desconocen tanto como el mejor modelo estocástico que se adecua a la misma, como los componentes de nivel, tendencia y estacionalidad de la función de recaudación tributaria. Utilizando el método de criterios de información y las técnicas de suavización exponencial, buscamos caracterizar, estimar y pronosticar. Observamos que la distribución de probabilidad que mejor ajusta es la log-normal. Además, el modelo de Holt Winters arroja los mejores resultados. Estos datos demuestran que tanto la distribución log-normal

1 El artículo es original producto de la investigación científica.

2 Licenciado en Economía egresado de la Facultad de Ciencias Económicas de la Universidad Nacional de la Plata (La Plata, Buenos Aires, República Argentina), Doctorando (en estado de tesis) de la carrera "Doctor en Ciencias Económicas" de la Universidad Nacional de La Matanza (San Justo, Buenos Aires, República Argentina). Profesor Adjunto en las materias "Economía" y "Probabilidad y Estadística" y "Metodología de la Investigación y Análisis Estadístico aplicado a la investigación científica" para las carreras de "Ingeniería Industrial", "Ingeniería en Energía" e "Ingeniería Electromecánica" en la Universidad Tecnológica Nacional Facultad Regional Santa Cruz (Río Gallegos, Santa Cruz, República Argentina). Investigador científico y tecnológico UTN. Categoría "D". Ordenanza N° 1341. Carrera Docente Investigador UTN.

ORCID: <https://orcid.org/0000-0003-3725-7451>

Email: [ariasmontes1973@gmail.com](mailto:ariasmontes1973@gmail.com)

Email: [juan.arias@frsc.utn.edu.ar](mailto:juan.arias@frsc.utn.edu.ar)

como el modelo aditivo de Holt Winters, son métodos que representan una adecuada caracterización, estimación y pronóstico de la recaudación de ingresos brutos, para la Provincia de Santa Cruz, República Argentina. Por último, los resultados de los métodos parecen coincidir con las estimaciones del organismo planificador, Ministerio de Economía, encontrándose la estimación puntual en un valor aproximado de 63 mil millones de pesos argentinos y el arrojado por el modelo de Holt Winters cercano a los 59 mil millones de pesos argentinos; para el ejercicio 2023.

**Palabras Claves:** Recaudación de impuestos de los ingresos brutos, Holt Winters aditivo, Suavizado exponencial, Pronósticos de serie de tiempo.

**Abstract:** Tax collection in subnational governments plays an important role in the process of budget formulation and strategic planning. Although the determining factors are known, they are unknown as well as the best stochastic model that fits it, as well as the components of: level, trend and seasonality of the tax collection function. Using the information criteria method and exponential smoothing techniques, we seek to characterize, estimate, and forecast. We observe that the probability distribution that best fits is the log-normal. Furthermore, the Holt Winters model yields the best results. These data show that both the log-normal distribution and the Holt Winters additive model are methods that represent an adequate characterization, estimation, and forecast of gross revenue collection for the Province of Santa Cruz, Argentine Republic. Finally, the results of the methods seem to coincide with the estimates of the planning body, the Ministry of Economy, finding the point estimate at an approximate value of 63 billion Argentine pesos and that produced by the Holt Winters model close to 59 million. million Argentine pesos; for the financial year 2023.

**Key words:** Tax collection of gross receipts, holt winters additive, exponential smoothing, Time series forecasts.

## 1. Introducción

El análisis de la recaudación tributaria en los gobiernos subnacionales desempeña un papel pivotal tanto en la formulación presupuestaria, ya sea en la micropresupuestación como en la macropresupuestación (Cornejo Rallo, 2022); como también en la planificación estratégica. Cornejo Rallo (2008) señala el carácter intuitivo de reconocer la centralidad del presupuesto para la gestión dentro de las organizaciones. Las agencias estatales necesitan recursos financieros para llevar adelante las tareas que les son encomendadas; el presupuesto determina cuales son los recursos disponibles durante un período establecido y el ámbito en el cual se utilizan. Ahora bien, la relación entre los presupuestos y la planificación estratégica presenta diversos problemas. Estos fueron redirigidos a la discusión entre formulación del presupuesto y la ejecución de las estrategias. Según Neilson (2008) la mejor ejecución de las estrategias se logra bajo 4 ejes centrales: derechos de decisión, flujo de información, cambios en las estructuras y el alineamiento con el presupuesto. Una parte central del trabajo se orienta a trabajar en la estimación de los recursos tributarios para lograr que la estrategia sea un proceso continuo, como lo señala Kaplan y Norton (2005); en concordancia con los flujos de información ascendentes y descendentes dentro de la organización gubernamental. Una estimación de recursos de carácter rigurosa y metodológicamente contrastable, contribuye de manera significativa a una mejor implementación de las estrategias. En uno de los procedimientos escalonados planteados por Kaplan y Norton (2005), se detallan los escalones. En primer lugar, se definen los objetivos e indicadores, en segundo lugar se asocia el objetivo con una fecha temporal y, en tercer lugar, se identifican las estrategias y los recursos necesarios para lograr los objetivos. Es en este último escalón es donde adquiere relevancia la estimación de los recursos.

Por otro lado, las causas de la recaudación tributaria han sido ampliamente investigadas (García, 2015; Martín, 2006). Dichos autores, hacen hincapié en las relaciones causales entre la variable a explicar, la recaudación tributaria y sus determinantes. Entre ellos se destacan las variables externas, el Producto Bruto Interno (PBI), Producto Bruto Geográfico (PBG), inflación, población económicamente activa, etc. Mientras que las internas van desde aquellas que se encuentran en la normativa misma (alícuota, base imponible), hasta las agrupadas en la gestión (intimaciones, soportes tecnológicos). En tanto Martín (2016) clasifica los determinantes en directos (legislación, materia gravada, normas de incumplimiento en el pago de obligaciones, etc.), y

aquellos denominados de segundo grado (necesidades de política fiscal, preferencias de los realizadores de política económica). En los últimos años ha aumentado el interés por explorar técnicas econométricas y probabilísticas, para avanzar tanto en la determinación de las formas funcionales como los parámetros y pronósticos. En un estudio exhaustivo de la recaudación tributaria en los gobiernos subnacionales, García (2015) encontró un modelo de regresión donde tanto el PBI como la población son estadísticamente significativas. A partir de los estudios anteriores, se proponen una estrategia nueva para trabajar en los pronósticos de la recaudación de ingresos tributarios, utilizando técnicas econométricas, llamadas técnicas de suavización exponencial. A pesar de su seguridad y eficacia, los estudios causales enfrentan varios inconvenientes importantes; entre ellos, el más importante es el acceso a la información, originando los problemas de especificidad en la relación de variables. Quizá, no tanto para los determinantes de segundo grado (agregados macroeconómicos), pero sí para los determinantes directos, para los cuales es sumamente limitado. Se han hecho esfuerzos para superar este problema mediante la utilización de diversas técnicas estadísticas. Las mismas van desde los modelos de series de tiempo hasta las técnicas de minería de datos que son supervisadas (árboles de decisión y árboles de regresión) y no supervisadas (reglas de asociación, clustering y redes neuronales artificiales). Sin embargo, continúa el debate sobre las mejores estrategias para la determinación de las características que explican la variable de interés, denominada recaudación tributaria. Hay muy poca investigación publicada sobre la determinación de la recaudación tributaria, en términos de su caracterización, estimación y pronóstico. Hasta la fecha, no han sido examinados exhaustivamente los diferentes parámetros de la estimación y sus implicancias. Hay muy pocos resultados publicados sobre la aplicación de técnicas de serie de tiempo en el estudio de la recaudación tributaria en gobiernos subnacionales para la República Argentina. Ninguna de estas cuestiones se exploró en los resultados presentados en García (2015) y Martín (2006).

El primer y principal reto es que las estimaciones econométricas sean el puntapié inicial para establecer controles de seguimiento y fortalecer la gestión pública conectando las estrategias con la estimación de recursos y, así, lograr los objetivos planteados (Cornejo Rallo, 2008). Como señala Escobar (2007) la utilización de las series de tiempo como herramienta tiene en cuenta que dado un conjunto de datos sobre un mismo fenómeno, existe una clara dependencia de los valores pasados, con lo cual conociendo datos del pasado se puede pronosticar el comportamiento de los datos en el futuro. En la misma línea, Collantes

(2004) señala tanto la facilidad en el desarrollo del modelo, como así también en la implementación sin dejar de lado la gran ductilidad en la interpretación y comprensión.

Un pronóstico de los ingresos tributarios resulta crucial para poder lograr los objetivos de política fiscal del gobierno. Los más relevantes, según Núñez Miñana (1998), los podemos resumir en: corrección en la asignación de recursos, satisfacción de las necesidades sociales a través de la provisión de bienes públicos y la redistribución del ingreso. Extensas investigaciones han destacado la importancia del análisis particular, correspondiente a la serie de ingresos brutos. La misma responde a que dicho tributo constituye porcentualmente el más relevante dentro del conjunto de tributos y donde cada gobierno subnacional define la base imponible; como señala Artana (2010).

Asimismo, la literatura no ofrece una metodología única para el abordaje de una mejor caracterización, estimación y pronóstico de la recaudación. Estas tres cuestiones actualmente no tienen respuesta y, por ello, se exploran aquí desde un punto de vista probabilístico y estadístico matemático. La necesidad surge debido a que los métodos actuales necesitan de un flujo de información que no se encuentra disponible. Estos estudios garantizan la complementariedad y una mejor comprensión de la recaudación en los gobiernos provinciales. Hay tres objetivos principales de este estudio: 1. Caracterizar. 2. Estimar y 3. Pronosticar, los componentes de la recaudación de ingresos brutos en la Provincia de Santa Cruz. Este estudio empleó la metodología de criterios de información (Queralt, 2015), para la selección del mejor modelo, junto con las técnicas de suavización exponencial triple (Winters, 1960, p.324), para la estimación y el pronóstico. En el mismo se ha identificado que la distribución de probabilidad que mejor se ajusta es la log-normal utilizando, además, como métricas los estadísticos de Anderson-Darling y Kolmogorov-Smirnov. Asimismo, la alta precisión medida a través del coeficiente de determinación  $R^2$  y el bajo RMSE (raíz del error cuadrático medio), permitieron discriminar el modelo Holt Winters como el más adecuado para la estimación y pronóstico. Específicamente, el modelo aditivo de Holt Winters posee nivel, tendencia y estacionalidad. Cada uno de dichos componentes son significativos estadísticamente. La serie de ingresos brutos fue obtenida del organismo recaudador subnacional denominado Agencia Santacruzense de Ingresos Públicos (ASIP), desde el período enero 2014 a noviembre 2022<sup>3</sup>. Los resultados obtenidos confirman la idoneidad del método para la estimación puntual,

---

3 Véase en <https://www.asip.gob.ar/recaudacion/>

correspondiente a los ingresos brutos. No se observaron diferencias entre la estimación expuesta por el Ministerio de Economía, Finanzas e Infraestructura de la Provincia de Santa Cruz para el ejercicio 2023 de aproximadamente 63 MM de pesos argentinos, contra la obtenida a través del modelo de Holt Winters de 59 MM pesos argentinos.

El presente estudio se organizó de la siguiente manera. En primer lugar, una introducción general sobre la revisión de la literatura y el planteamiento del problema. Luego, en segundo lugar, los métodos y materiales utilizados, detallando los procedimientos para el logro de los objetivos, incisos 2.1) Caracterización, 2.2) Estimación y Pronóstico 2.3) Precisión y 2.4) Esquema de combinación. En tercer lugar, se presentan los resultados obtenidos sobre 3.1) Caracterización, 3.2) Estimación de los parámetros, 3.2.1) Tendencia, 3.2.2) Factores Estacionales, 3.2.3) Esquema aditivo e Independencia, 3.2.4) Error, 3.2.5) Outliers y 3.3) Pronósticos. En el apartado cuarto, hemos pretendido poner de manifiesto la discusión sobre la problemática bajo estudio. Mientras que en el apartado quinto de conclusiones se revela el grado de precisión para lograr los objetivos del estudio. Finalizando, con el apartado sexto de referencias bibliográficas.

## 2. Método y Materiales

### 2.1 Caracterización

Estudios anteriores, Queralt (2015), han basado sus criterios de selección de modelos a través de los criterios de información (IC). Hay dos tipos de diseño que se utilizan: AIC (Akaike Information Criterion) y BIC (Bayesian Information Criterion).

Cuando tenemos  $k$  modelos alternativos, denominados  $M_1, M_2, M_3, \dots, M_k$ , siendo  $k$  el número de parámetros que tiene el modelo  $M_k$ . Vamos a elegir aquel  $k$  que minimice la siguiente ecuación:

$$IC(k) = \ln \sigma_k^2 + k C(T)/T$$

Donde:

$\sigma_k^2$  = Varianza residual del Modelo  $M_k$

$T$  = Tamaño muestral

$C(T)$  = Término de penalización

Para el criterio AIC tenemos que  $C(T) = 2$ , mientras que en BIC

el término de penalización es  $\ln T$ . Como afirma Martínez (2009), el resultado final es aquel modelo que tiene el AIC más bajo.

Por otro lado y siguiendo a Cabeza (2007), también se utilizó tanto el estadístico de Anderson-Darling (A-D), en el cual se va asignando un peso diferente a las discrepancias entre las distribución empírica y teórica. Mientras que el de Kolmogorov-Smirnov, permite trabajar con el grado de acuerdo entre las distribuciones empírica y teórica.

## 2.2 Estimación y pronóstico

En un artículo reciente, Chung Pinzás (2023), sostiene que los estudios a través del Modelo Holt Winters (H-W) ofrecen una gran capacidad a los cambios en los componentes de una serie de tiempo. Sin embargo, este método no es válido para analizar las tendencias a largo plazo, el mismo requiere calibraciones periódicas.

Este método es particularmente útil para estudiar las componentes de una serie de tiempo. En este aspecto, puede ser suficiente en muchos casos prácticos para obtener una forma funcional y realizar pronósticos.

El estudio utilizó una muestra de conveniencia con datos publicados por la Agencia Santacruceña de Ingresos Públicos (ASIP), ente recaudador de la Provincia de Santa Cruz. Dicha serie corresponde a los Ingresos Brutos (IIBB) durante el periodo enero 2014 a noviembre 2022. El método de recolección de los datos es retrospectivo.

Las ecuaciones utilizadas para el modelo, se tomaron de la literatura (García Zare, 2015, p.99)

1. Serie Exponencial

$$A_t = \alpha Y_t / E_t - L + (1 - \alpha) (A_{t-1} + T_{t-1})$$

2. Componente Tendencia

$$T_t = \gamma (A_t - A_{t-1}) + (1 - \gamma) T_{t-1}$$

3. Componente Estacionalidad

$$E_t = \delta Y_t / A_t + (1 - \delta) E_{t-L}$$

4. Pronóstico

$$Y_{t+p} = (A_t - p T_t) E_{t-L+p}$$

Donde:

- $A_t$  = nuevo valor atenuado
- $\alpha$  = constante de nivel ( $0 < \alpha < 1$ )
- $Y_t$  = nueva observación o valor real de la serie en el periodo  $t$
- $\gamma$  = constante de atenuación de estimación de la tendencia ( $0 < \gamma < 1$ )

- $T_t$  = estimación de la tendencia
- $\delta$  = constante de atenuación de estimación de la estacionalidad ( $0 < \delta < 1$ )
- $S_t$  = estimación de la estacionalidad
- $p$  = periodos a estimar a futuro
- $L$  = longitud de la estacionalidad
- $Y_{t+p}$  = pronóstico de  $p$  periodos a futuro

Este método requiere un gran número de pruebas estadísticas y su pronóstico depende de tener un modelo especificado correctamente y con validez interna.

### 2.3 Precisión y bondad

En primer lugar, una de las métricas que vamos a utilizar es el RMSE (raíz del error cuadrático medio), para poder evaluar el grado de ajuste entre las alternativas de modelado, del conjunto de datos. La fórmula empleada para realizar dicho procedimiento es:

$$RMSE = \sqrt{[\sum (P_i - O_i)^2 / n]}$$

Donde:

$\sqrt{\quad}$  = Raíz Cuadrada

$\sum$  = sumatoria

$P_i$  = es el valor pronosticado para la  $i$ -ésima observación en el conjunto de datos

$O_i$  = es el valor observado para la  $i$ -ésima observación en el conjunto de datos

$n$  = es el tamaño de la muestra

Cuanto mayor sea el RMSE, mayor será la diferencia entre los valores pronosticados y los valores observados, lo que significa que peor se ajusta el modelo de regresión a los datos. En cambio, cuanto más pequeño sea el RMSE, mejor podrá un modelo ajustar los datos observados. Resulta sumamente útil comparar el RMSE de dos modelos diferentes entre sí para ver qué modelo se ajusta mejor.

En segundo lugar, el MAPE (error porcentual absoluto medio) arroja un porcentaje del error. Con lo cual resulta un indicador de fácil lectura e interpretación. En términos de la medición de la exactitud, valores pequeños indican un ajuste adecuado. El MAPE está expresado de la siguiente manera:

$$MAPE = 100/N \times \sum |(x - xf)/x|$$

Donde:

$x$  = son las observaciones actuales de las series de tiempo

$x_f$  = son las series de tiempo estimadas o pronosticadas

$N$  = es el número de observaciones

Por último, para la bondad del ajuste utilizamos el coeficiente de determinación,  $R^2$  siempre está entre 0 y 100%. En los casos en que arroje el valor de 0% indica que el modelo no explica ninguna porción de la variabilidad de los datos de respuesta en torno a su media. Mientras que el 100% indica que el modelo explica toda la variabilidad de los datos de respuesta en torno a la media del conjunto de datos (Pere Grima et al., 2011, p. 238). El  $R^2$ , se calcula de la siguiente manera:

$$R^2 = SSR / SST = 1 - SSE / SST$$

$$SST = \sum (x - x_{media})^2$$

$$SSR = \sum (x' - x_{media})^2$$

$$SSE = \sum (x - x_{media})^2$$

Donde:

SST= Corresponde a la suma de cuadrados total

SSR= Corresponde a la suma de cuadrados de regresión

SSE= Corresponde a la suma de cuadrados de error

## 2.4 Esquema de combinación

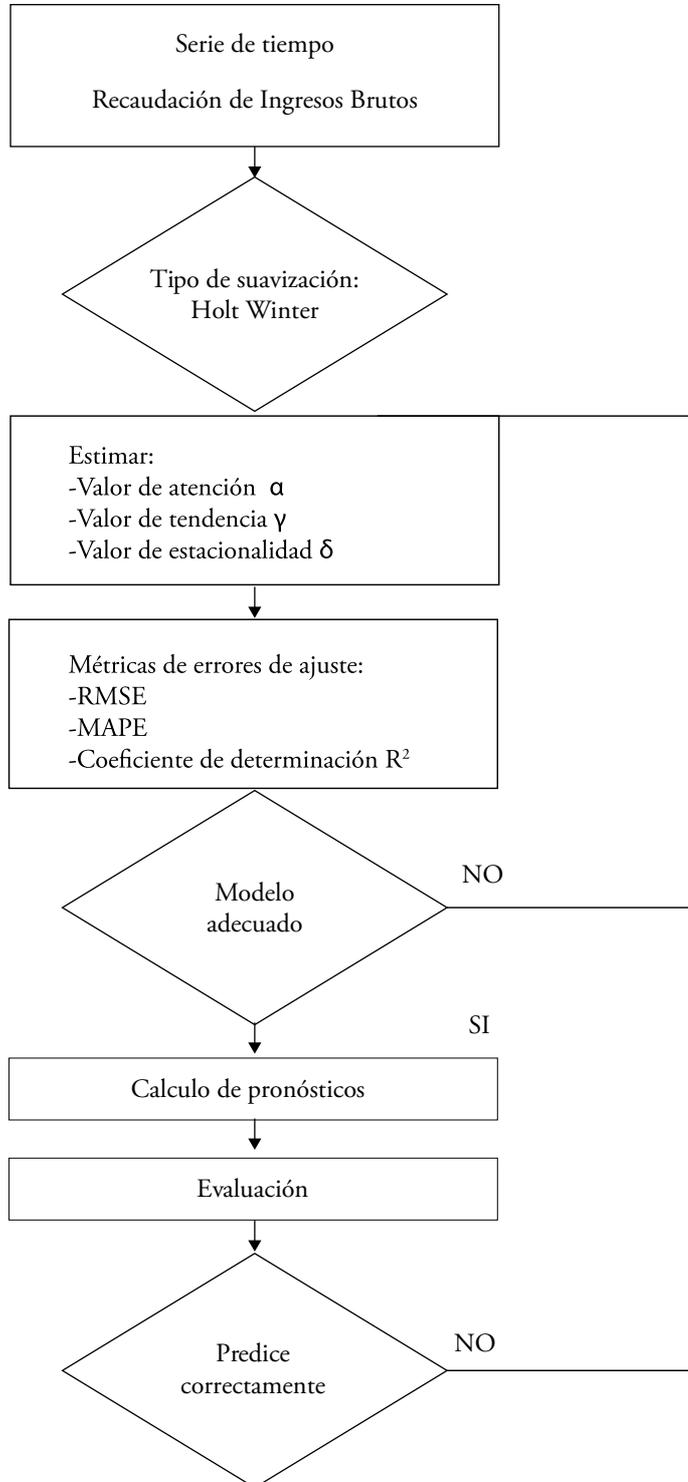
La técnica de las diferencias ayuda a determinar el tipo de esquema en el cual se combinan los componentes en las series temporales (Quezada, 2015, p. 260). La misma surge de comparar los coeficientes de variación de Pearson para las series diferencias ( $CVP_d$ ) y aquellos denominados cocientes estacionales ( $CVP_c$ ). El esquema será aditivo o multiplicativo de acuerdo a la siguiente regla de decisión:

Si el  $CVP_c > CVP_d$ , tenemos un esquema aditivo.

Si el  $CVP_c < CVP_d$ , tenemos un esquema multiplicativo.

Por último, para la caracterización de los datos, estimación y pronóstico se trabajó con datos secundarios (estudios retrospectivos) y se utilizaron herramientas de procesamiento estadístico, con los paquetes estadísticos IBM SPSS Statistics V22.0 y Risk simulator.

Gráfico 1. Esquema de Trabajo



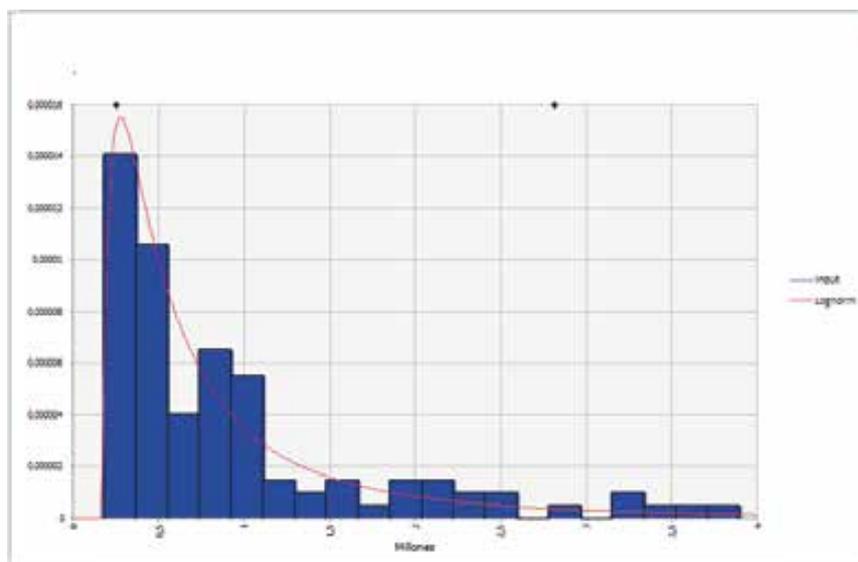
### 3. Resultados

Con el fin de caracterizar, estimar y pronosticar los ingresos brutos en la Provincia de Santa Cruz, se procedió a la aplicación de técnicas de suavización exponencial junto con los criterios de información. A continuación, se utilizó un conjunto de pruebas estadísticas para encontrar la distribución, que se ajusta a los datos. Asimismo, la comparación de los resultados revela una forma funcional triple: tendencia, estacionalidad y ruido.

#### 3.1 Caracterización

Del Gráfico 2 se desprende que el conjunto de datos se asemeja a una Distribución de probabilidad log-normal, conforme al criterio de selección AIC, encontrándose claramente inclinada hacia la derecha del cuadrante. La misma es utilizada para modelar fenómenos que toman valores no negativos. La observación más significativa de este estudio es que dicha distribución refleja el comportamiento de los datos. La probabilidad de que sea cero es extremadamente baja y los valores más probables están por encima de cero.

Gráfico 2. Distribución log-normal Ingresos Brutos, Provincia de Santa Cruz



Fuente: Elaboración propia a partir de los datos de la ASIP.

En la siguiente “Tabla 1” se pueden observar los estadísticos más relevantes de la distribución de probabilidad log-normal, para la serie de datos ingresos brutos:

Tabla 1. Estadísticos Distribución log-normal

Distribución log-normal	
Mínimo	156,7844
Máximo	+Infinity
Media	984,7502
Moda	274,2088
Mediana	588,563
Desviación estándar	1354,6988
Asimetría	9,2887

Fuente: Elaboración propia, a partir de los datos de la ASIP.

En la misma, observamos que la esperanza matemática o la media (984,75) en la distribución log-normal es mayor que su mediana (588,56). De esta manera se le da más importancia a los valores más altos que una distribución normal para los percentiles 5% y 50%.

Tabla 2. Estadísticos Anderson-Darling (A-D) y Kolmogorov-Smirnov (K-S)

	Gauss Inversa	Log-normal	LogLogística	Exponencial	Triangular	Logística	Normal	Uniforme
A-D	0,42	<b>0,39</b>	0,52	1,16	Infinito	5,89	8,0	65,2
p-valor	(n/d)	(n/d)	(n/d)	(n/d)	(n/d)	(n/d)	(n/d)	(n/d)
K-S	0,07	<b>0,06</b>	0,07	0,09	0,34	0,19	0,19	0,51
p-valor	(n/d)	(n/d)	(n/d)	(n/d)	(n/d)	(n/d)	(n/d)	(n/d)

n/d: no disponible

Fuente: Elaboración propia, a partir de los datos de la ASIP.

Tanto el estadístico de Anderson-Darling como el de Kolmogorov-Smirnov, nos indican qué tanto se ajustan los datos de una serie de tiempo a una distribución específica. Para este conjunto de datos se optó por compararlas con las distribuciones: Gauss Inversa, log-normal, Log logística, Exponencial, Triangular, Normal y Uniforme. A partir de la comparación de este subconjunto de distribuciones, se obtiene en la

distribución log-normal los estadísticos más bajos, tanto en el criterio de A-D como en el de K-S, siendo para el primero de 0,39 y 0,06 para el segundo, como se observa en la “Tabla 2”. Tanto los valores de p-valor, para las pruebas de Anderson-Darling, como las de Kolmogorov-Smirnov, no están disponibles porque no existen matemáticamente para estos casos. A sabiendas que el mejor ajuste de la distribución de datos la encontramos en un menor valor del estadístico. Dicho esto, concluimos que el mejor ajuste para este conjunto de datos lo encontramos en la distribución de probabilidad log-normal.

### 3.2 Estimación de los parámetros

Para encontrar el modelo que mejor se ajusta a los datos, se realizaron las estimaciones entre los diferentes modelos de suavización exponencial y los modelos ARIMA (Modelo autorregresivo integrado de media móvil). Utilizando como indicador para la selección el RMSE (raíz del error cuadrático medio).

Tabla 3. Ajustes del modelo suavización exponencial Holt Winters

Estadístico de ajuste	Media	Mínimo	Máximo
R cuadrado	0,991	0,991	0,991
<b>RMSE</b>	<b>79,604</b>	<b>79,604</b>	<b>79,604</b>
<b>MAPE</b>	<b>6,385</b>	6,385	6,385
MaxAPE	31,908	31,908	31,908
MAE	53,490	53,490	53,490
MaxAE	286,333	286,333	286,333
BIC normalizado	8,885	8,885	8,885

Fuente: Elaboración propia, a partir de los datos de la ASIP.

Tabla 4. Ajustes del modelo ARIMA (0,1,0) (0,1,0)

Estadístico de ajuste	Media	Mínimo	Máximo
R cuadrado	0,978	0,978	0,978
<b>RMSE</b>	<b>127,944</b>	<b>127,944</b>	<b>127,944</b>
<b>MAPE</b>	<b>8,697</b>	8,697	8,697
MaxAPE	44,385	44,385	44,385
MAE	83,706	83,706	83,706
MaxAE	512,315	512,315	512,315
BIC normalizado	9,752	9,752	9,752

Fuente: Elaboración propia, a partir de los datos de la ASIP.

El aspecto más interesante de las “Tablas 3” y “Tabla 4” precedentes, radica en los valores del estadístico de ajuste RMSE (Raíz del error cuadrático medio), y MAPE y  $R^2$ . Y tal como muestra la tabla que arroja los resultados del conjunto de los modelos de suavización exponencial. El mejor ajuste se obtiene con estos modelos y dentro de ellos el aditivo de Holt Winters que tiene nivel, tendencia y estacionalidad.

Tabla 5. Parámetros del modelo suavizado aditivo de Holt Winters

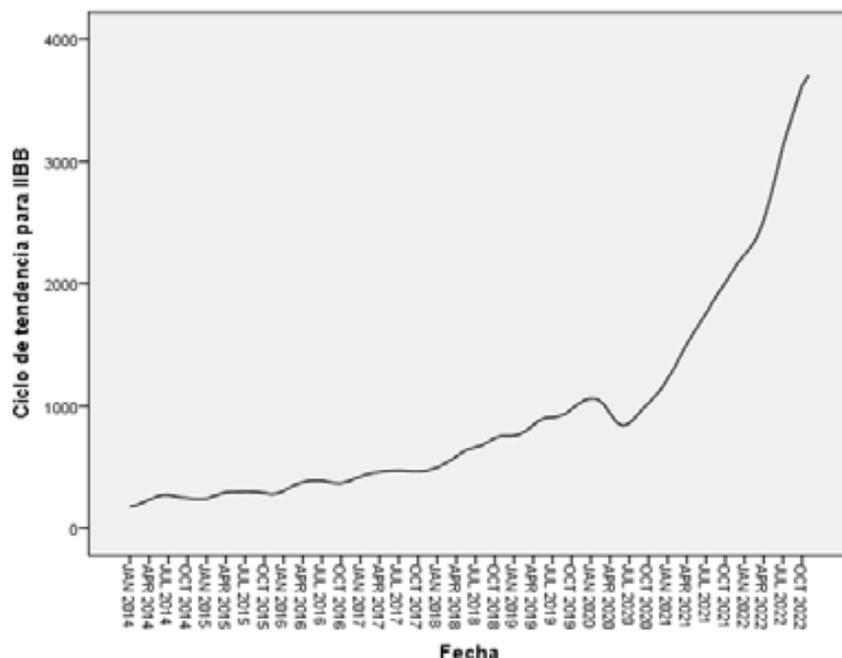
Modelo		Estimación	Desvío estándar	t	p-valor	
Ingresos Brutos	Sin transformación	$\alpha$ (nivel)	0,665	0,089	7,509	0,000
		$\gamma$ (tendencia)	0,204	0,073	2,782	0,006
		$\delta$ (estacionalidad)	0,768	0,293	2,620	0,010

Fuente: Elaboración propia, a partir de los datos de la ASIP.

La “Tabla 5” es bastante reveladora en varios sentidos. En primer lugar, a diferencia de las otras tablas tenemos desagregados los componentes de la serie de tiempo. Según estas estimaciones, el modelo que mejor se ajusta es el aditivo de Holt Winters. Este es lineal, arrojando un valor inicial de nivel ( $\alpha$ ) de 0,665, tendencia ( $\gamma$ ) positiva de 0,204 y un componente estacional  $\delta = 0,768$ , aditivo de periodos anual ( $p=12$ ). En segundo lugar, dichos componentes resultan significativos estadísticamente,  $\alpha$  ( $p=0,000$ ),  $\gamma$  ( $p=0,006$ ) y  $\delta$  ( $p=0,010$ ),  $R^2=0,991$ .

### 3.2.1 Tendencia

Gráfico 3. Comportamiento de la tendencia



Fuente: Elaboración propia, a partir de los datos de la ASIP.

Esta serie temporal contiene datos, correspondientes al ciclo de tendencia, sobre la recaudación de ingresos brutos en la Provincia de Santa Cruz, República Argentina, desde enero 2014 a noviembre 2022. En el mismo se observa la representación gráfica de esta serie de tiempo, donde los valores se encuentran en millones de pesos argentinos. Asimismo, la serie de valores de tendencia está mucho más alisada y sugiere una tendencia lineal creciente.

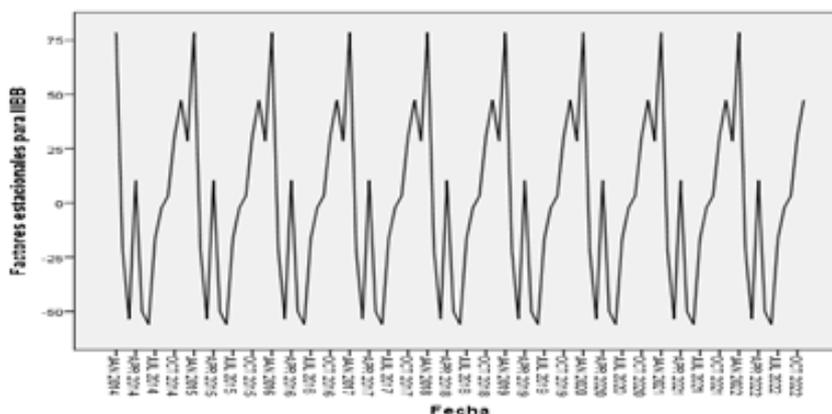
### 3.2.2 Factores Estacionales

Tabla 6. Factores Estacionales

Período	Factor estacional
1-Enero	78,589
2-Febrero	-21,604
3-Marzo	-53,359
4-Abril	10,481
5-Mayo	-50,018
6-Junio	-55,951
7-Julio	-16,324
8-Agosto	-2,059
9-Septiembre	3,255
10-Octubre	30,814
11-Noviembre	47,481
12-Diciembre	28,695

Fuente: Elaboración propia, a partir de los datos de la ASIP.

Gráfico 4. Descomposición Estacional



Fuente: Elaboración propia a partir de los datos de la ASIP.

La “Tabla 6” y el “Gráfico 4” revelan que la serie en los meses:

enero, abril, septiembre, octubre, noviembre y diciembre toman valores superiores a la tendencia media; el componente estacional tiene mayor repercusión, incrementando la recaudación de dicho tributo. Asimismo, en los meses de febrero, marzo, mayo, junio, julio y agosto es cuando se produce el mayor decremento de los ingresos debido a la estacionalidad. Por consiguiente, la variación de la serie de datos en los diferentes meses del año es sumamente similar todos los años, lo cual nos indicaría claramente, que existe estacionalidad<sup>4</sup>.

### 3.2.3. Esquema Aditivo e Independencia

Para el análisis de la determinación del esquema respectivo, utilizamos el método analítico de diferencias y encontramos que el coeficiente de variación de *Pearson* para los cocientes ( $CVP_c$ ) ha sido de 10,60 y el coeficiente de variación de *Pearson* de diferencias ( $CVP_d$ ) fue de 1,10. Siendo, por lo tanto, el esquema aditivo. Finalmente, mediante la función Box test se aplicó la prueba de Ljung-Box a la serie para verificar la independencia de datos. Los resultados se muestran a continuación.

$H_0$ : Los datos de la serie se distribuyen de forma independiente

$H_1$ : Los datos de la serie no se distribuyen de forma independiente

X-squared = 1,733, df = 15, p-valor = 0,277

En la “Tabla 7”, se observa que p es mayor que 0,05, por lo tanto se acepta  $H_0$ , es decir, los datos de la serie se distribuyen independientemente.

Tabla 7. Estadísticos para determinar la independencia del modelo

Modelo	Número de predictores	Estadísticos de ajuste del modelo	Ljung-Box Q			Número de valores atípicos
		R cuadrado estacionaria	Estadísticos	DF	p-valor	
Ingresos Brutos (\$Millones)	0	0,357	17,733	15	0,277	0

Fuente: Elaboración propia, a partir de los datos de la ASIP.

4 Con el método de la media móvil, en primer lugar se obtiene el componente ciclo más tendencia y luego los índices de valores estacionales (IVE). Luego, se realiza la ponderación móvil, en este caso tomamos aquella que dice ponderación por 5, dado que el periodo es par e igual a 12 meses.

### 3.2.4. Error

Para conocer si la descomposición estacional es correcta, necesitamos que la media del error (residuo) sea igual a cero, dado que encontramos que el mejor modelo es el aditivo de Holt Winters. Realizamos la siguiente prueba de hipótesis para el residuo:

Siendo el modelo aditivo de Holt Winters.

$$X_t = T_t + E_t + R_t$$

Donde:

$X_t$  = el valor observado de la serie temporal en el periodo t

$T_t$  = el componente tendencia en el periodo t

$E_t$  = el componente estacional en el periodo t

$R_t$  = el componente residuo en el periodo t

Prueba de hipótesis para el residuo ( )

Ho: Media del residuo = 0

H1: Media del residuo  $\neq$  0

p -valor= 0,882

Por lo tanto, concluimos que los datos de los residuos tienen una media igual a cero evaluado a través del test de la t de student,  $t(106) = -0,149$ ,  $p\text{-valor} > 0,05$  (IC 95 % -8,54 a 7,34). Encontrando, que la descomposición estacional está correctamente especificada (“Tabla 8”).

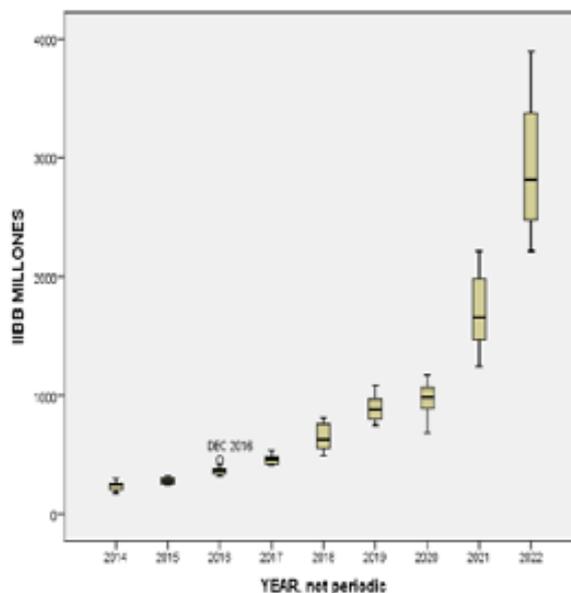
Tabla 8. Prueba t de student para una muestra única

	Valor de prueba = 0					
	t de student	Grados de libertad	p-valor	Diferencia de medias	95% de intervalo de confianza de la diferencia	
					Inferior	Superior
Error para Ingresos Brutos	-0,149	106	0,882	-0,596	-8,54	7,34

Fuente: Elaboración propia, a partir de los datos de la ASIP.

### 3.2.5. Valores Atípicos (Outliers)

Gráfico 5. Diagrama de Caja de Bigotes



Fuente: Elaboración propia, a partir de los datos de la ASIP.

Los valores atípicos generan conflictos durante el proceso de análisis de los mismos, lo que resulta en información inconsistente (Orellana, 2020,p.57), por consiguiente, la identificación de los mismos resulta central, dado que afectan la capacidad de pronóstico del modelo de serie temporal aditivo de Holt Winters. Tal como se observa en el “Gráfico 5”, vemos que la serie no presentó valores atípicos.

### 3.3. Pronóstico

Utilizando el modelo que en la sección 3.2.3 y 3.2.4 ha sido validado, se procedió a realizar los pronósticos para el período 2023.

Tabla 9. Pronóstico de los Ingresos Brutos, para la Provincia de Santa Cruz, para el período diciembre 2022 a diciembre 2023, con intervalos de predicción al 95%

Año	Mes	Valor Pronosticado	Límite Inferior	Límite superior
2022	12	3.945	3.787	4.103
2023	1	4.184	3.981	4.386
2023	2	4.161	3.910	4.411
2023	3	4.324	4.022	4.626
2023	4	4.616	4.259	4.973
2023	5	4.663	4.248	5.079
2023	6	4.857	4.380	5.333
2023	7	5.126	4.586	5.667
2023	8	5.232	4.625	5.839
2023	9	5.440	4.764	6.116
2023	10	5.546	4.798	6.293
2023	11	5.812	4.991	6.634
2023	12	5.877	4.963	6.792

Fuente: Elaboración propia, a partir de los datos de la ASIP.

Los resultados demuestran la idoneidad del método para realizar un pronóstico anual. En conjunto, estos resultados muestran que el bajo error y la alta precisión permiten seleccionar, al modelo aditivo de Holt Winters como la mejor selección entre todos los de suavización exponencial y los autoregresivos integrados de media móviles.

#### 4. Discusión

El presente estudio estimó los componentes de la ecuación que muestran nivel, tendencia y estacionalidad para la serie de ingresos brutos de la Provincia de Santa Cruz. Estos resultados corroboran y amplían los hallazgos de gran parte de los trabajos anteriores sobre estimación y pronósticos a través de modelos de suavización exponencial para series de tiempo (Avenidaño et al., 2021). En línea con el trabajo de Iglesias (1970), este enfoque pone énfasis en la medición y el pronóstico de las formas de financiamiento de los gobiernos subnacionales en la República Argentina. Esto puede considerarse un avance significativo en términos de dotar a

las agencias gubernamentales de herramientas de planificación, control y auditoría de los recursos tributarios; junto con la mejora de la eficiencia y productividad en la realización de las tareas propias del organismo (Cornejo Rallo, 2008, p. 113). La posibilidad de tener una metodología de estimación de ingresos, robusta y verificable otorga la posibilidad de trabajar con presupuestos operativos creíbles y alcanzables. La agencia gubernamental encargada de la elaboración del presupuesto provincial en la Provincia de Santa Cruz, es decir, el Ministerio de Economía, Finanzas e Infraestructura (Cornejo Rallo, 2022, p. 129) en su elevación del proyecto de presupuesto, que luego fue aprobado por la legislatura por Ley N° 3812/2023, estimó para el período 2023 la suma de \$ 63.810 en miles de millones de pesos argentinos. En tanto las estimaciones, junto con el modelo aditivo de Holt Winters, arrojaron un valor cercano a \$59.838 mil millones, si tomamos el valor puntual de la estimación. La brecha entre cada uno de los pronósticos es de aproximadamente el 7%.

Además, un punto a favor de esta investigación es el hecho de que aborda el nivel, la tendencia y la descomposición estacional de la serie de tiempo. Encontrando una caída en los meses invernales de la recaudación. A pesar de la novedad de este hallazgo, estos resultados se encuentran limitados por razones metodológicas de tamaño de muestra y falta de estudios previos.

Otro hallazgo importante en la presente investigación es la determinación de la función de distribución de probabilidad log-normal para la serie datos e ingresos brutos. En la misma línea, un estudio reciente de Fernández (2023) desarrolla una distribución de probabilidad, con la serie de recaudación tributaria como porcentaje del producto para analizar la probabilidad de ocurrencia de diferentes escenarios planteados a priori, con la finalidad de la construcción de un fondo anti-cíclico.

A pesar de que nuestros resultados coinciden en algunos aspectos con estudios anteriores (Villegas, 2014), nuestro hallazgo difiere de la mayoría de los otros estudios debido a la utilización de las técnicas de suavización exponencial de Holt Winters. Además, por lo que sabemos, este es el más sólido estudio publicado hasta la fecha sobre estimación y pronóstico de ingresos tributarios para gobiernos subnacionales en la República Argentina. Sin embargo, reconocemos que nuestros resultados pueden estar limitados por la naturaleza intrínseca del estudio. La mayoría de los estudios paramétricos (métodos de regresión) y no paramétricos, como las redes neuronales artificiales (RNA), son de tipo causal y arrojan muchas más información en términos de relación de variables. Sin embargo, las mismas no se encuentran exentas de limitaciones. Como señala Escobar (2007), el principal inconveniente en los modelos RNA es

la elección de cada elemento en la arquitectura de red cuando los métodos de regresión se enfrentan a los problemas de especificidad (Iglesias, 1979) en las relaciones funcionales de las variables. Esto último se contrapone con los de serie de tiempo, en términos de complejidad y facilidad de interpretación.

## 5. Conclusiones

En conclusión, nuestros resultados revelan que la precisión en el pronóstico de la serie ingresos brutos, utilizando el modelo de suavización triple de Holt Winters, alcanzó un RMSE (Raíz del error cuadrático medio) de 79,60; siendo muy por debajo de los que arrojó el modelo auto regresivo integrado de media móvil ARIMA (0, 1,0) (0, 1,0) con un valor de 197,94. Asimismo, el Error Porcentual Absoluto Media (MAPE), brinda un mejor desempeño del pronóstico en el modelo de suavización de HW con un valor del 6,38%, en tanto que el modelo auto regresivo es de 8,67%. Mientras que la precisión medida a través del coeficiente de determinación R<sup>2</sup>, alcanzó para el modelo de Holt Winters un índice del 99,1%; y en el modelo autorregresivo integrado de medias móviles ARIMA (0, 1,0) (0, 1,0) uno de 97,8 %. Por otro lado, la distribución log-normal es la que mejor se ajusta a los datos. Los estadísticos Anderson-Darling, como los de Kolmogorov-Smirnov obtenidos, han arrojado los valores más bajos en comparación con las otras distribuciones contrastadas, validando la hipótesis de que este conjunto de datos asemeja su comportamiento a una distribución log-normal.

Nuestros resultados apoyan la importancia de formalizar el comportamiento de la serie de datos a través de su caracterización, estimación y pronóstico de los parámetros. En consecuencia, esto proporciona una importante contribución en términos de poder discriminar el modelo más adecuado para ser utilizado en las estimaciones de ingresos brutos por parte de las autoridades gubernamentales. Sin embargo, se necesita más investigación para evaluar, contrastar y confirmar la precisión de los pronósticos.

## 6. Bibliografía

- Artana, D., & Templado, I. (2010). Is the Argentine revenue effort “too” high. Documento de Trabajo N°106. Fundación de Investigaciones Económicas Latinoamericanas. Recuperado de <http://www.fiel.org/publicaciones/Documentos/DOCTRAB106.pdf>
- Avendaño, S., & Hernández, Marta. (2021). Pronóstico del tipo de cambio USDMXN durante el COVID-19 con métodos de suavización y descomposición. *Revista Enfoques: Ciencia Política y Administración Pública*, 19(34), 1-16. Recuperado de <http://www.revistaenfoques.cl/index.php/revista-uno/article/view/552>.
- Cabeza, M., & Torra, S. (2007). El riesgo en la empresa. Medición y control mediante @Risk. USA: Palisade Corporation, 119-121.
- Collantes Duarte, Joanna. Colmenares La Cruz, Gerardo. Orlandoni Merli, Giampaolo. & Rivas Echeverría, Franklin. (2004). A comparison of time series forecasting between artificial neural networks and box and jenkins methods. *Revista Técnica de la Facultad de Ingeniería Universidad del Zulia*, 27(3), 146-160. Recuperado de [http://ve.scielo.org/scielo.php?script=sci\\_arttext&pid=S02547702004000300002&lng=es&tlng=es](http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S02547702004000300002&lng=es&tlng=es).
- Cornejo Rallo, M. (2008). Gestión Pública, implementación de estrategias y presupuesto. El caso del Servicio de Impuestos Internos chileno. *Revista Enfoques: Ciencia Política y Administración Pública*, 6(9), 103-125. Recuperado de <http://www.revistaenfoques.cl/index.php/revista-uno/article/view/204>.
- Cornejo Rallo, M. (2022). Incrementalismo presupuestario en el Gobierno Central chileno. Evidencia para el decenio 2012 - 2021. *Revista Enfoques: Ciencia Política Y Administración Pública*, 20(37), 127-162. Recuperado de <http://www.revistaenfoques.cl/index.php/revista-uno/article/view/609>.
- Chung Pinzás, A., & Inche Mitma, J. (2023). Aplicación de técnicas Econométricas y Machine Learning para Series de tiempo Univariadas en una escuela profesional de una universidad pública. *Veritas*, 23(1), 47-51. doi:10.35286/veritas.v23i1.335.
- Escobar, R., Valdez, H., Santiago, C. (2007) Redes Neuronales en predicción de series de tiempo. Una aplicación a la Industria. *Revista Ciencia y Tecnología*, 9(34), 33-47. Recuperado de <https://www.palermo.edu/ingenieria/investigaciondesarrollo/revistacienciaecnologia/edicion-9.html>.
- Fernández, M., & Malament, M. (2023). Hacia la construcción de un fondo fiscal anticíclico. Universidad del CEMA, (847). Recuperado de <https://econpapers.repec.org/paper/cemdoctra/847.htm>

- García, L., & Gutiérrez, M. (2015). Relevamiento de la situación tributaria de las unidades subnacionales en Argentina. Determinantes de la función de recaudación impositiva en Chaco, Chubut, Misiones, Salta, San Luis, Santa Cruz, Río Negro. *Revista Enfoques: Ciencia Política y Administración Pública*, 13(22), 111-134. Recuperado de <http://www.revistaenfoques.cl/index.php/revista-uno/article/view/334>.
- García Zare, E. & Bermúdez Edó, J. D. (2015). Análisis estadístico del modelo de Holt-Winters: Aplicación a la contaminación de aire por pm2.5 de Lima, Perú. *Conocimiento para el desarrollo*, 6(2). Recuperado a partir de <https://revista.usanpedro.edu.pe/index.php/CPD/article/view/84>.
- Iglesias, A. (1979). Documento de trabajo Análisis Econométrico de la Recaudación Tributaria en Chile: 1960-1976. Instituto de Economía Facultad de Economía y Administración UC,31(66). Recuperado de [https://economia.uc.cl/docs/dt\\_066.pdf](https://economia.uc.cl/docs/dt_066.pdf).
- Kaplan, R., & Norton, D. (2005). *La organización focalizada en la estrategia*. Barcelona: editorial Gestión 2000.
- Ley Provincial N°3.812. Presupuesto General de la Administración Pública Provincial para el año 2023. Provincia de Santa Cruz, República Argentina. 15 de Diciembre 2022. Decreto N°1602 /2022.
- Martin, Fernando. (2006). Los factores determinantes de la recaudación tributaria. Gobierno de Argentina [www.mecon.gov.ar/sip/basehome/factores\\_recaudacion\\_tributaria.pdf](http://www.mecon.gov.ar/sip/basehome/factores_recaudacion_tributaria.pdf).
- Martínez, D. R., Albin, J., Cabaleiro, J., Pena, T., Rivera, F., & Blanco, V. (2009). El Criterio de Información de Akaike en la Obtención de Modelos Estadísticos de Rendimiento. XX Jornadas de Paralelismo, 439-444. Recuperado de [https://scholar.google.com/scholar?cluster=5152783290322529646&hl=es&as\\_sdt=0,5](https://scholar.google.com/scholar?cluster=5152783290322529646&hl=es&as_sdt=0,5)
- Neilson, G., Martin, K., Powers, E. (2008). Los secretos de una ejecución exitosa de la estrategia. *Harvard Business Review*,86(6).
- Núñez, Miñana. (1998). *Finanzas Públicas*. Buenos Aires: Ediciones Macchi.
- Orellana, Marcos. & Cedillo, Priscila. (2020). Detección de valores atípicos con técnicas de minería de datos y métodos estadísticos. *Enfoque UTE*, 11(1), 56-67. <https://doi.org/10.29019/enfoque.v11n1.584>.
- Pere Grima, C., Lluís, A., & Tort-Martorell, Xavier. (2011). *Estadística con Minitab: Aplicaciones para el control y la mejora de la calidad*. Madrid: Ibergarceta Publicaciones S.L.
- Queralt, R. (2015). Simulación de series temporales: una aplicación al precio del petróleo. *Palisade Conferencias de Riesgo 2015. Mejoras Prácticas en riego y análisis de decisiones*. 19 de Mayo de 2015. Madrid.
- Quezada, Lucio. (2015). *Estadística con SPSS 22*. Perú: Editorial Macro.

- Sotelo, A. (2008). La relación planificación-presupuesto en el marco de la gestión orientada a resultados. *Revista del CLAD Reforma y Democracia*, (40).
- Villegas, P. (2014). Pronóstico de ingresos tributarios recaudados por impuestos generales a la venta a nivel nacional por la SUNAT, mediante la metodología BOX-JENKINS y redes neuronales artificiales para el año 2014 [tesis]. Universidad Nacional Pedro Gallo. Disponible en: <https://repositorio.unprg.edu.pe/handle/20.500.12893/459>
- Winters, P. R., (1960). Forecasting Sales by Exponentially Weighted Moving Average. *Management Sciences*, 6, 324-342.