



Análisis de implementación de métodos de *nowcasting* en el INEGI

Junio 2026



ANTECEDENTES

- En diversos foros internacionales y nacionales, una recomendación reiterada que se realiza a las Oficinas Nacionales de Estadística es que la información tenga las siguientes características:
 - **Oportuna, Frecuente, y Granular.**
- En relación con el primer aspecto —la **oportunidad**—, existe una amplia experiencia internacional en la implementación de métodos de *nowcasting* para estimar de manera temprana y precisa variables relevantes de la coyuntura económica.
- Entre los países que han desarrollado este tipo de prácticas se encuentran **Alemania, Australia, Belice, Canadá, El Salvador, Estados Unidos, Nueva Zelanda, Reino Unido y Suiza.**



ANTECEDENTES



- El INEGI ha acumulado experiencia en la elaboración de **tres indicadores** de *nowcasting*:
 - Indicador Mensual Oportuno de la Actividad Manufacturera (**IMOAM**), implementado experimentalmente desde agosto de 2015.
 - Indicador Oportuno de la Actividad Económica (**IOAE**), disponible desde octubre de 2020.
 - Indicador Oportuno del Consumo Privado (**IOCP**), publicado desde diciembre de 2022.
- Asimismo, cuenta con una **estimación oportuna** de **carácter oficial**:
 - Estimación Oportuna del Producto Interno Bruto (**EOPIBT**)



Nowcasting

- **Backcasting (Ejemplo: retropolación del PIB)**

- $E(y_{T-H}) = E[y_{T-H} | \Theta_X g(\mathbf{X}_{T-H}), \Theta_Y f(y_{T-H+1}, \dots, y_T)]$

- **Forecasting**

- $E(y_{T+H}) = E[y_{T+H} | \Theta_X g(\hat{\mathbf{X}}_{T+H}), \Theta_Y f(y_{T+H-1}, \dots, y_{T+H-p})]$

- **Nowcasting (Ejemplos: IMOAM, IOAE e IOCP)**

- $E(y_{T+H}) = E[y_{T+H} | \Theta_X g(\mathbf{X}_{T+H}), \Theta_Y f(y_{T+H-1}, \dots, y_{T+H-p})]$

- **Flash estimate (Ejemplo: EOPIBT)**

- $E(y_{T+H}) = E[y_{T+H} | \Theta_X g(\mathbf{X}_{T+H}), \Theta_Y f(\tilde{y}_{T+H})]$



OBJETIVOS

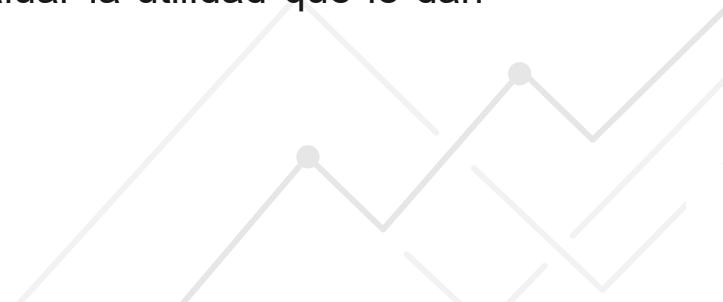


- **Objetivo general**

- Evaluar la pertinencia de aplicar métodos de *nowcasting* en el INEGI, considerando los criterios de precisión del error de estimación, relevancia para la toma de decisiones y robustez estadística de los resultados.

- **Objetivos específicos**

- **Evaluar históricamente** el desempeño de los métodos de *nowcasting* implementados por el INEGI.
- **Construir criterios empíricos** que permitan determinar umbrales aceptables de error de estimación.
- **Realizar pruebas de robustez estadística** que garanticen la estabilidad y confiabilidad de los resultados.
- **Incorporar análisis de sentimientos** como fuente complementaria para evaluar la utilidad que le dan los usuarios.



IOAE e IOCP – SÍNTESIS METODOLÓGICA



$$Y_t = PF_t + \varepsilon_t$$

$$F_t = \Phi F_{t-1} + \eta_t$$

$$\varepsilon_t = \Gamma \varepsilon_{t-1} + a_t$$

IOAE

IOCP

- Transformaciones de correlación máxima contemporánea.
- Selección de tópicos de *Google Trends*.
- Estimación del factor común: PC + Filtro de Kalman (en tiempo real).
- Validación cruzada para la selección de modelos.
- Evaluación de resultados vs competidores en tiempo *pseudo-real*.

Horizonte T+1

- Uso exclusivo de series estacionarias.
- Aplicación del criterio de causalidad de Granger y de transformaciones de correlación máxima para la selección de predictores.
- Estimación de modelos basados en enfoques de aprendizaje automático, MIDAS y factores dinámicos.
- Ensamble de modelos.
- Validación cruzada para la selección y evaluación del modelo final.

Horizonte T+2

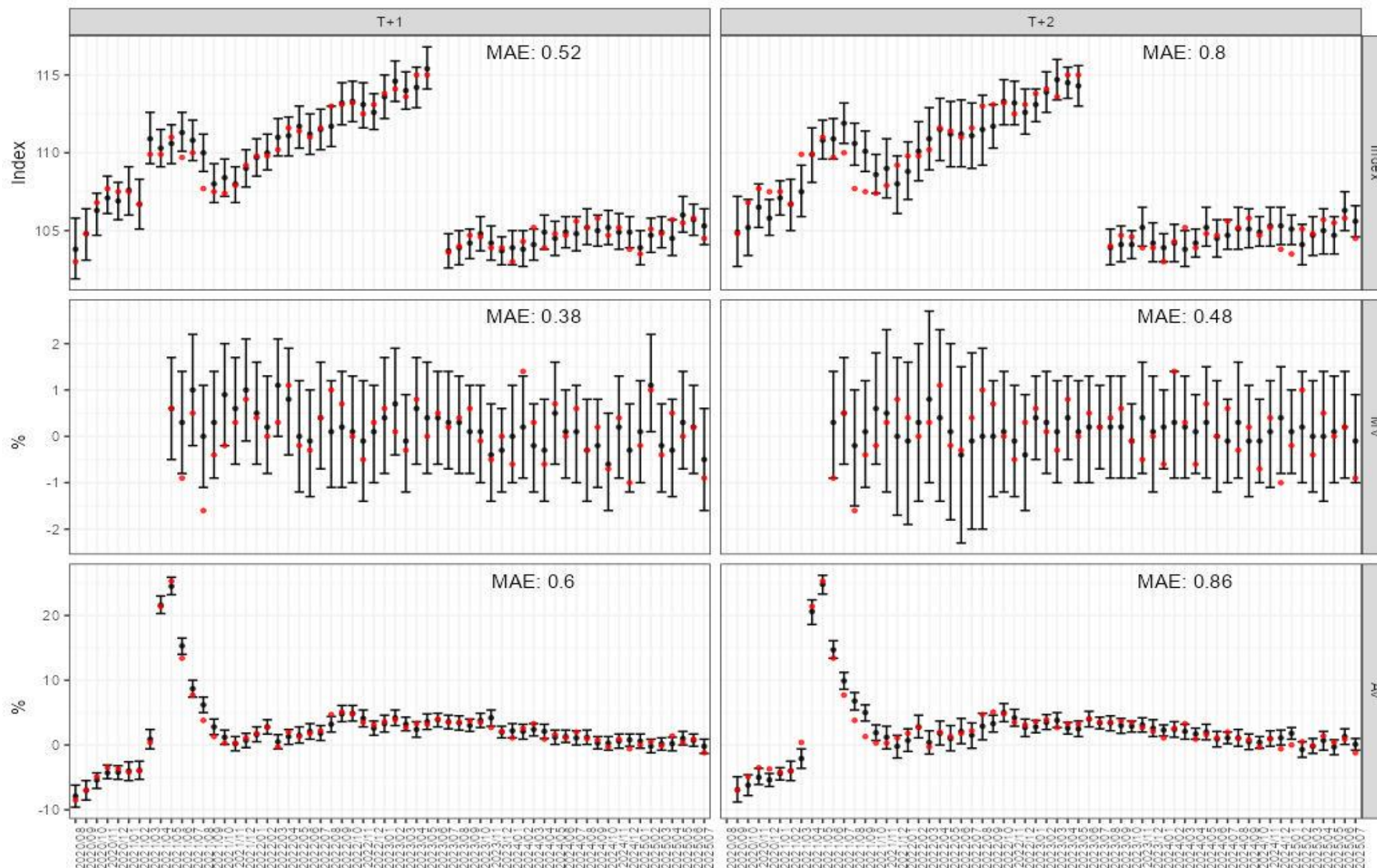
- Se asume el valor estimado de T+1 como conocido.
- Se aplica la metodología del IOAE para generar el *nowcast* correspondiente a T+2.

Resultados históricos

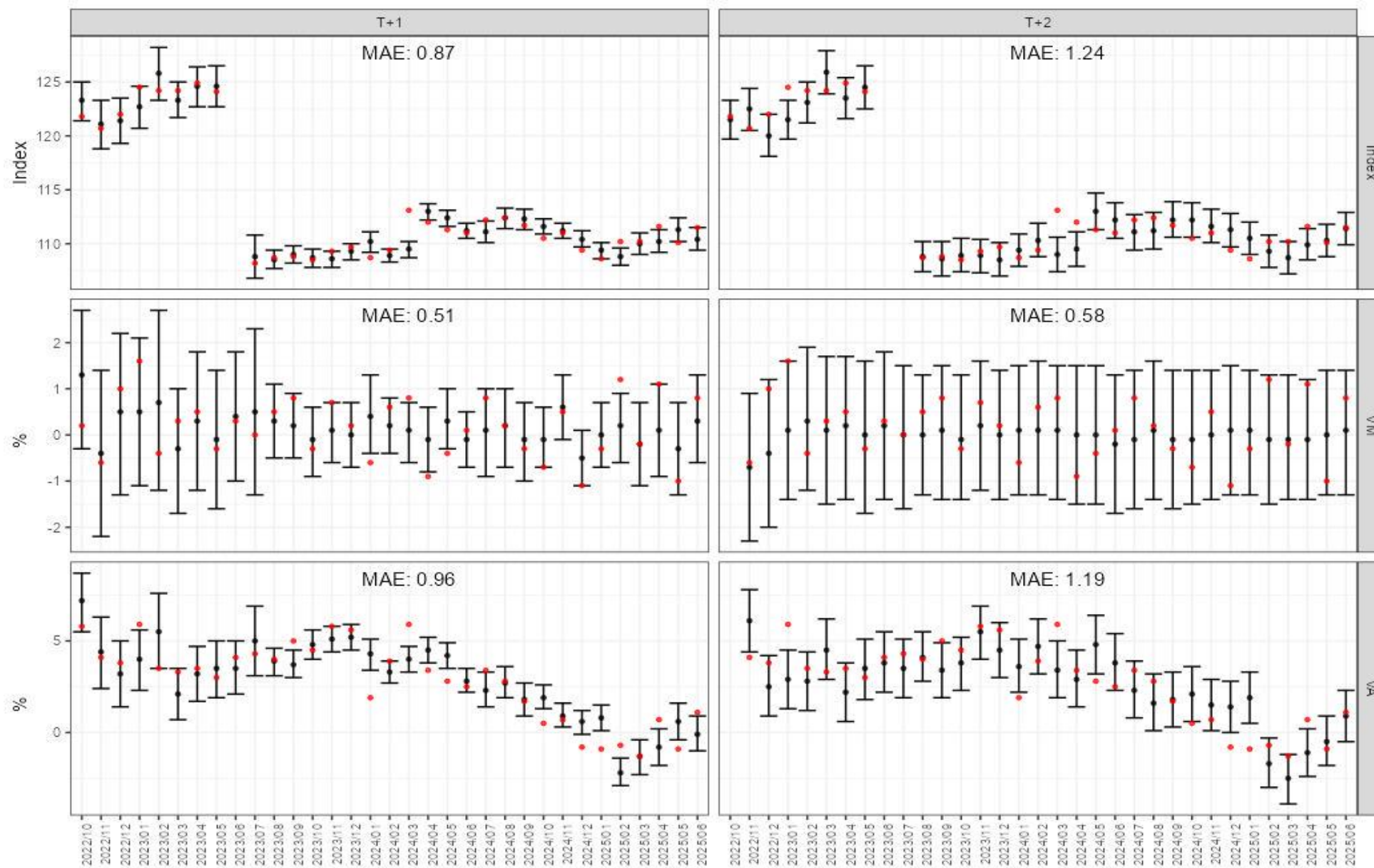
IOAE e IOCP



IOAE
Error
histórico



IOCP Error histórico



ANÁLISIS DE SENSIBILIDAD

- El objetivo es modelar la relación entre las características del entorno de estimación y el error resultante del *nowcast*.

$$E_i = f(\mathbf{Z}_i) + e_i,$$

donde

E_i es el error de estimación relativo (MSE, RMSE, MAE, MAPE, etc.),

\mathbf{Z}_i es vector de condiciones (días de publicación, tipo de indicador, si es usado para estadística oficial, tipo de error, interacciones, etc.),

$f(\cdot)$ se aproxima mediante regresión penalizada (LASSO), lo que te permite obtener una función de respuesta parsimoniosa.



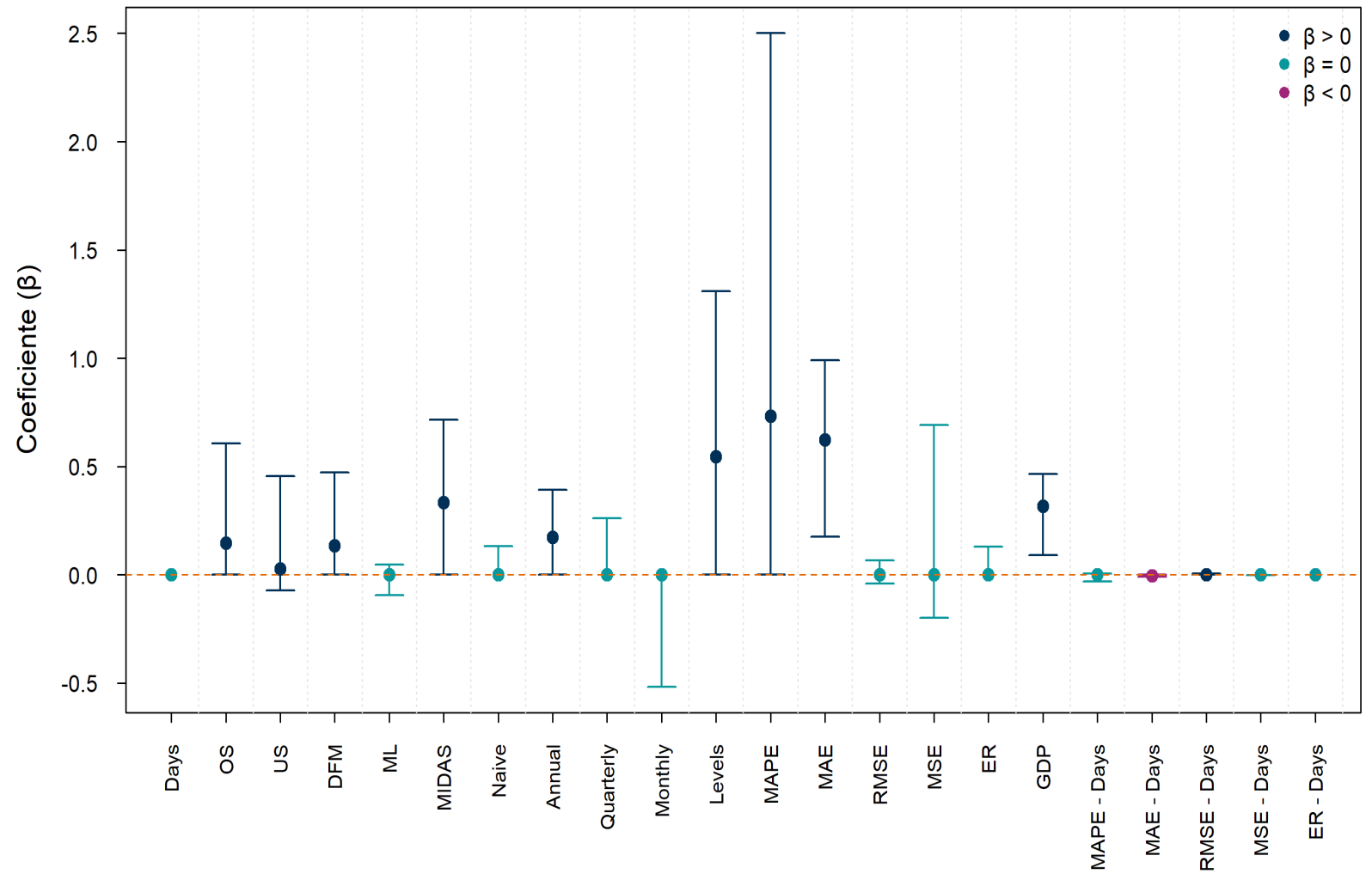
DEFINICIÓN DE UMBRALES

- Se aplicó un *Bootstrap* sobre las predicciones del modelo LASSO para cuantificar la incertidumbre asociada a los errores esperados.
- El proceso consiste en:
 1. Re-muestrear con reemplazo las observaciones del conjunto original de entrenamiento (miles de veces, p. ej. 25,000 iteraciones).
 2. En cada iteración, reentrenar el modelo LASSO y generar nuevas predicciones.
 3. Con los resultados de todas las iteraciones, se construye una distribución empírica del error esperado.
 4. A partir de dichas simulaciones se obtienen intervalos de confianza para los coeficientes de regresión (95%), así como los cuantiles Q1, Q2 y Q3, que permiten identificar zonas de menor o mayor riesgo de error.

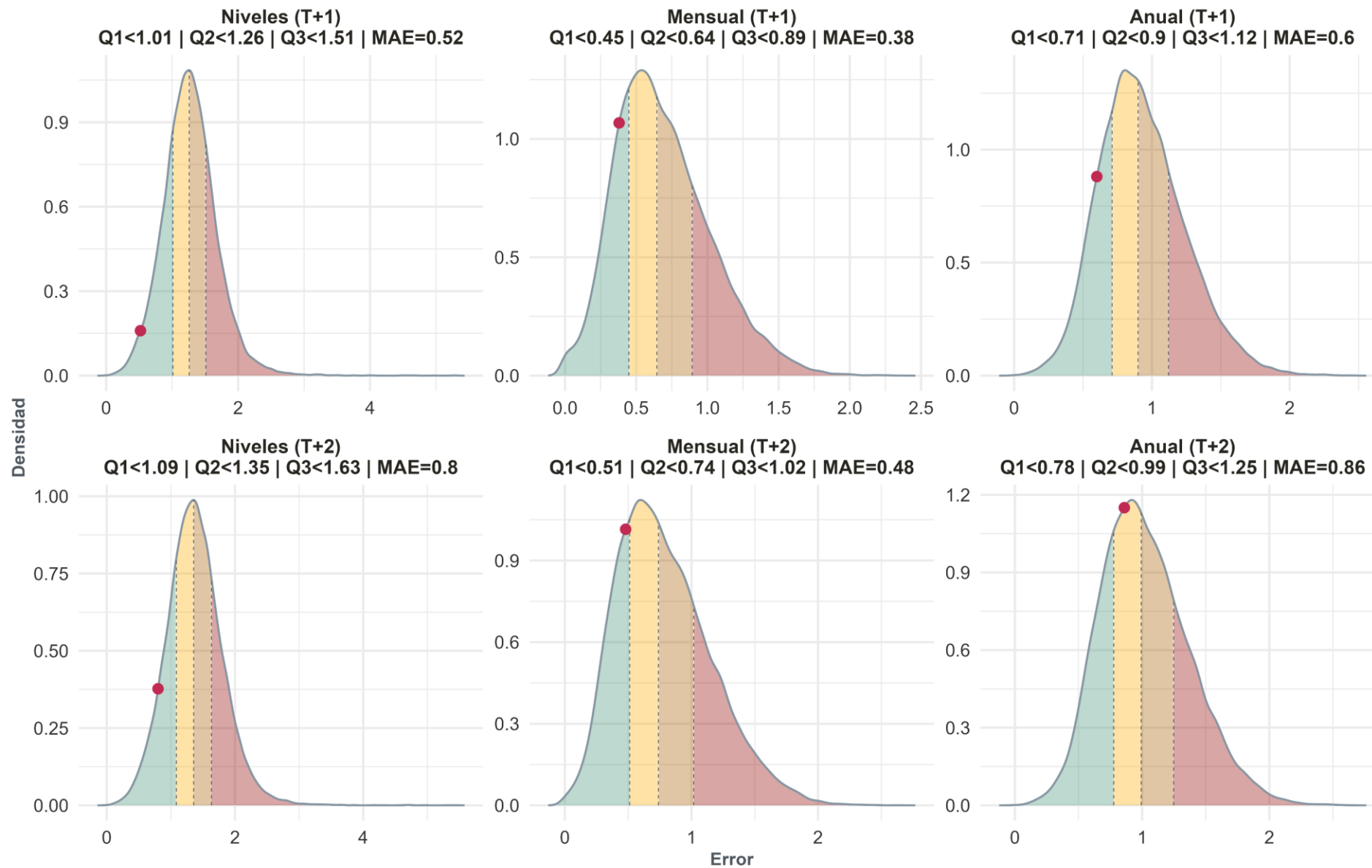
Intervalos de confianza estimados vía Bootstrap: Regresión LASSO

IOAE

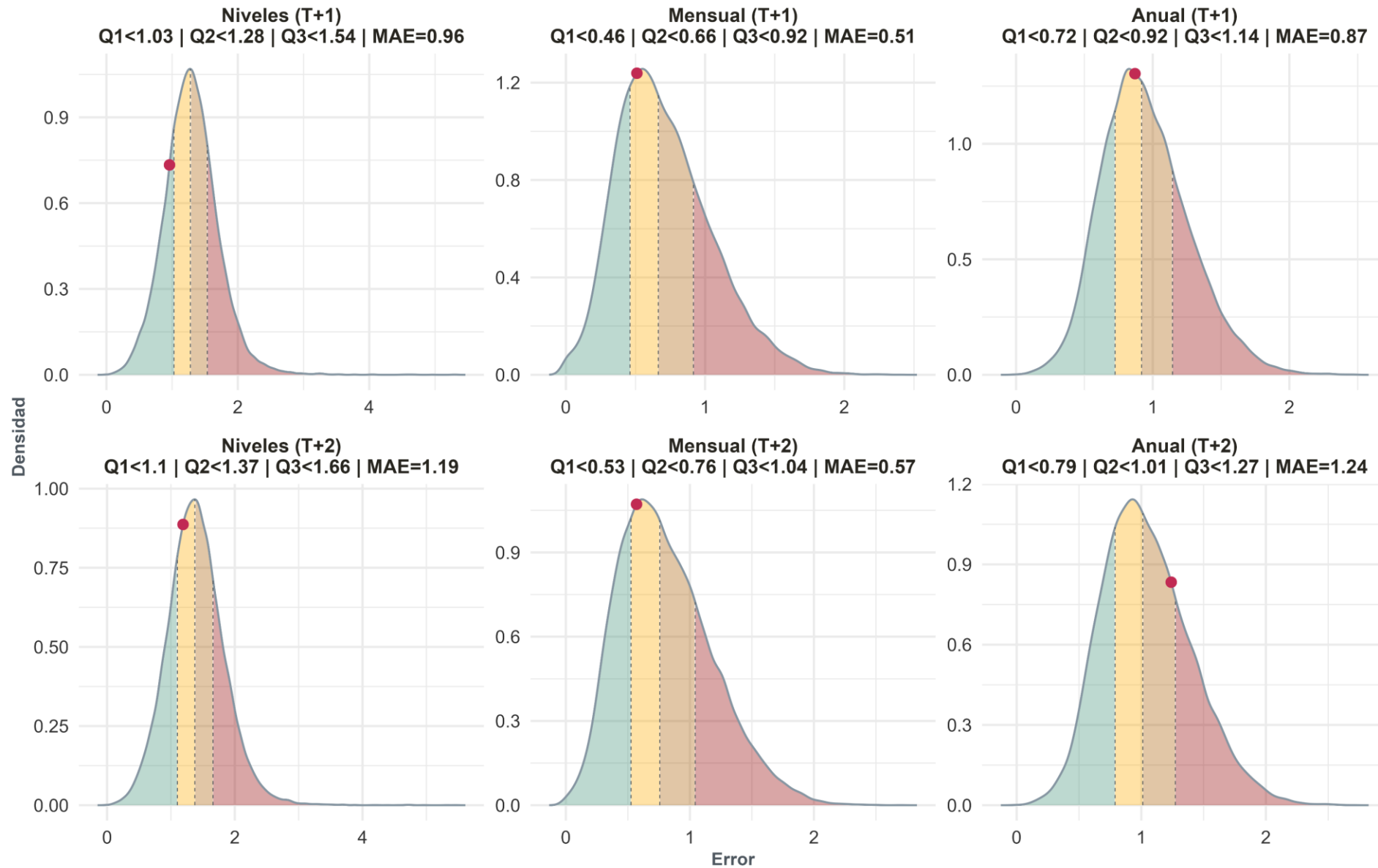
Coeficientes estimados del modelo LASSO



IOAE — Densidad del error esperado por escenario (Bootstrap)



IOCP — Densidad del error esperado por escenario (Bootstrap)



Análisis de sentimientos X: ejemplificación IOAE



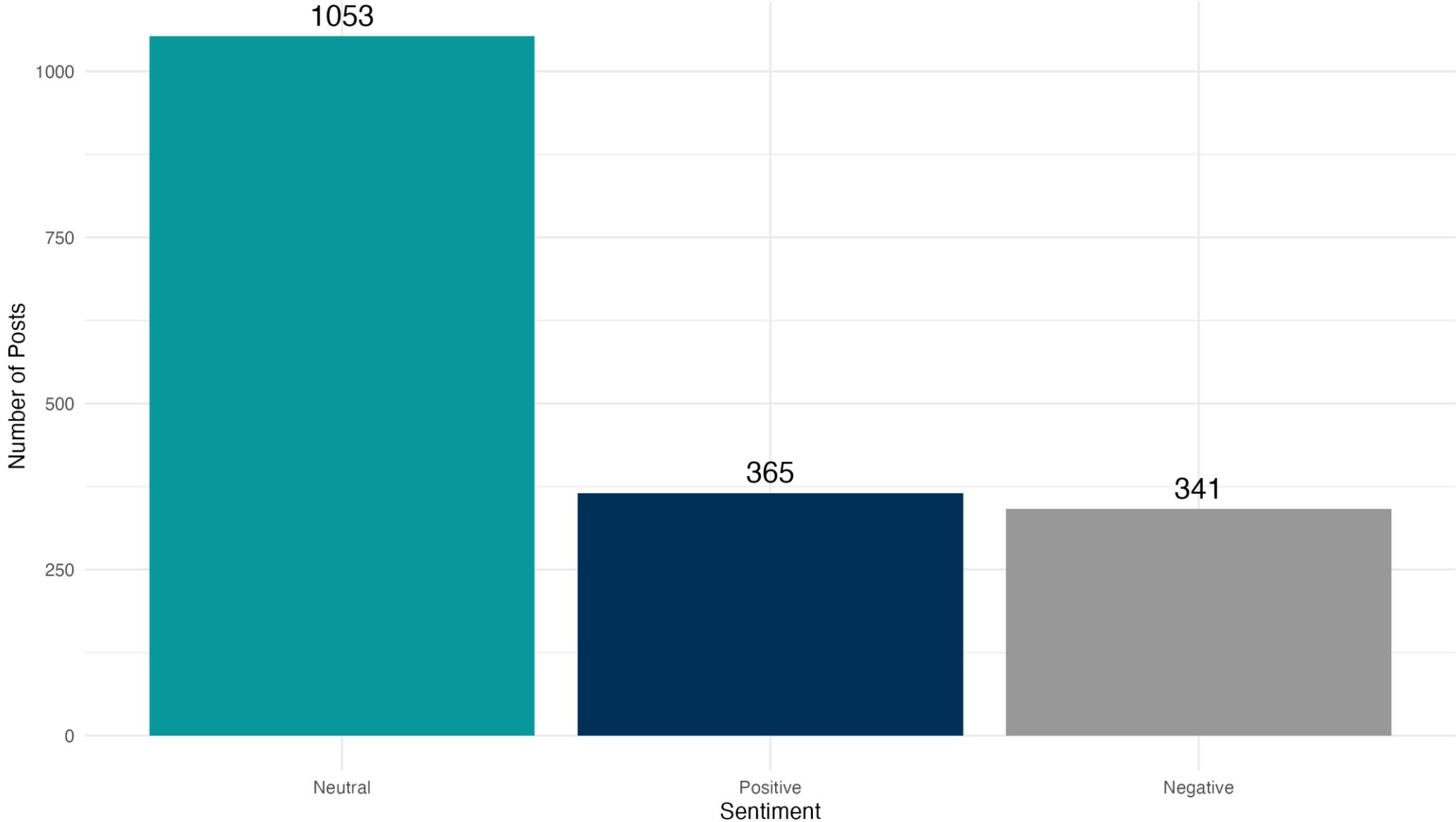
ANÁLISIS DE SENTIMIENTOS X: IOAE



- Se aplicaron técnicas de Procesamiento de Lenguaje Natural para analizar publicaciones en X (antes Twitter) relacionadas con el IOAE.
- Se utilizó el modelo BETO (BERT en español) para clasificar el sentimiento de cada tuit como positivo, neutral o negativo.
- El proceso incluyó: tokenización del texto, predicción del modelo y transformación softmax para determinar la probabilidad de cada clase.
- Las categorías se definieron como:
 - Positivo: opiniones favorables u optimistas sobre el IOAE.
 - Neutral: mensajes informativos o descriptivos sin carga emocional.
 - Negativo: comentarios críticos o pesimistas.



Sentiment Distribution in X (Twitter)



Análisis de robustez: ejemplificación IOAE



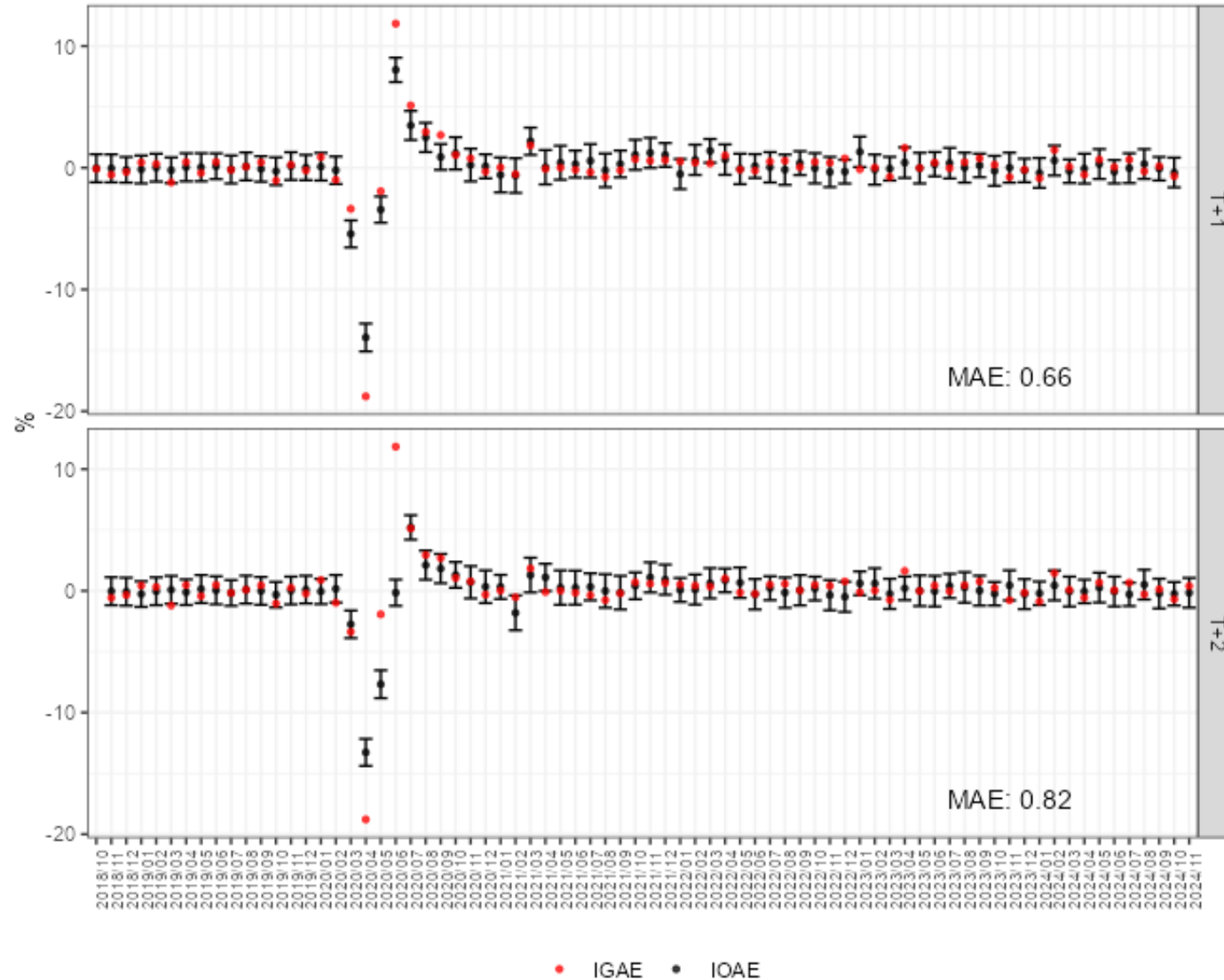
ANÁLISIS DE ROBUSTEZ IOAE

- Se realizaron estimaciones en tiempo *pseudo-real* para los horizontes **T+1** (octubre de 2018 a 2024) y **T+2** (noviembre de 2018 a 2024), aplicando la **metodología del IOAE** y, adicionalmente, los siguientes enfoques comparativos:
 - **Regresiones univariadas:** estimaciones oportunas del IGAE utilizando individualmente cada una de las series disponibles para T+2.
 - **Regresión LASSO:** estimación del IGAE dos pasos hacia adelante mediante validación cruzada con ventana expansiva, para seleccionar las variables más relevantes y controlar el sobreajuste.
 - **FAVAR:** modelos VAR ampliados con factores comunes.
 - **MLP:** red neuronal multicapa extendida con factores latentes como entradas adicionales.
 - **ARIMA:** modelo puramente univariado de series de tiempo.

ANÁLISIS DE ROBUSTEZ IOAE

- Se estimaron las **pruebas de Diebold–Mariano** con corrección **HAC** para evaluar la **superioridad estadística** del **IOAE** frente a sus respectivos modelos comparativos.
- Se implementaron pruebas Superior Predictive Ability (**SPA**, Hansen, 2005) para concluir estadística si algún competidor del IOAE tiene un rendimiento superior.
- Adicionalmente, se aplicó la **prueba *Model Confidence Set* (MCS)** propuesta por **Hansen et al. (2011)**, con el fin de realizar comparaciones múltiples entre modelos.
- Las funciones de pérdida empleadas fueron el **Error Absoluto Medio (MAE)** y la **Raíz del Error Cuadrático Medio (RMSE)**.
- Asimismo, se realizaron estimaciones adicionales considerando:
 - Diferentes **números de factores comunes ($r = 1,2,3$)**.
 - **Esquemas de validación** con ventanas **expansivas** y **móviles**.
 - **Distintos periodos muestrales**, incluyendo la **muestra completa** y una versión **excluyendo el periodo asociado a la COVID-19**.

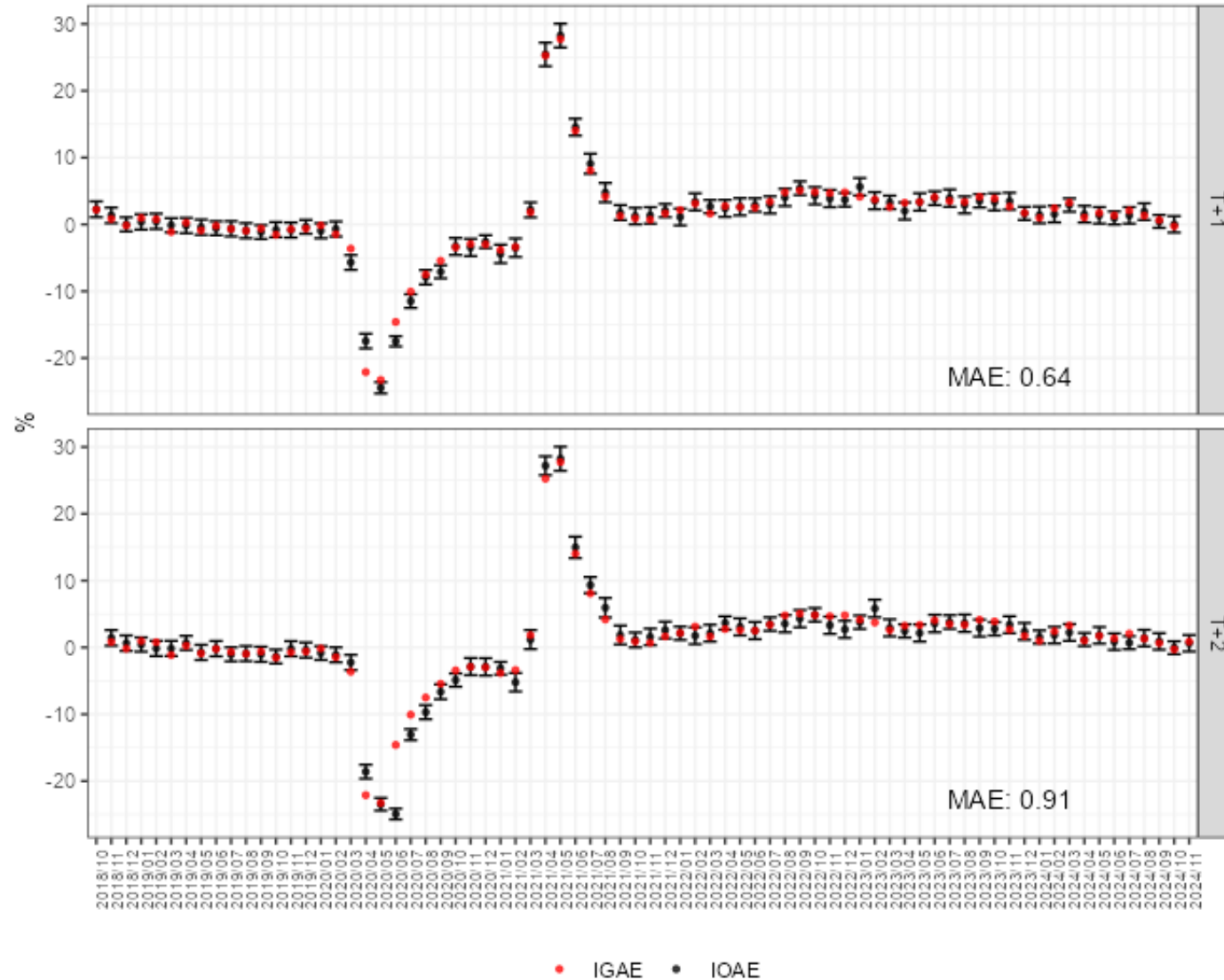
Estimación en tiempo pseudo-real IOAE en variaciones mensuales



Errores de **0.38** y **0.48** en tiempo real, para T+1 y T+2 respectivamente.

Estimaciones con cifras desde mayo 2021.

Estimación en tiempo pseudo-real IOAE en variaciones anuales



Errores de **0.60** y **0.86** en tiempo real, para T+1 y T+2 respectivamente.

Estimaciones con cifras desde agosto 2020.

RESULTADOS DE COMPARACIONES

- **DM-HAC**

- La prueba **DM-HAC** muestra que el **IOAE** supera estadísticamente (al 10%) a todas las regresiones individuales y al resto de modelos, salvo la **regresión LASSO en T+1**, con la que es equivalente, aunque con ligera ventaja para LASSO.

- **SPA**

- Ningún modelo supera al IOAE ni en **T+1** ni en **T+2** bajo las dos funciones de pérdida.

- **MCS**

- Utilizando el **MAE**, los resultados obtenidos mediante la **prueba MCS** son consistentes con los de las pruebas **DM-HAC**.
- Al emplear el **RMSE**, los **indicadores de Estados Unidos** —particularmente el **desempleo** y la **actividad manufacturera**— muestran también un **buen desempeño para el horizonte T+2**.
- Finalmente, al considerar el **MASE** (*Mean Absolute Scaled Error*), las conclusiones se mantienen equivalentes a las obtenidas con el MAE, es decir, **a favor de la metodología IOAE**.

INTERPRETACIÓN DE RESULTADOS

- El IOAE presenta los **mejores resultados de predicción**, particularmente para el horizonte **T+2**, y un **desempeño competitivo para T+1** en comparación con la **regresión LASSO**.
 - El IOAE **sintetiza información contemporánea**.
 - La regresión LASSO **aprovecha información disponible únicamente para T+1**.
 - La **introducción de no linealidades** o de **rezagos adicionales del IGAE no mejora** el desempeño de las estimaciones oportunas.
 - Los **modelos puramente univariados de series de tiempo** muestran un **rendimiento inferior**.



ROBUSTEZ

- En la **muestra completa**, el uso de **tres factores comunes** reduce el error de estimación para el horizonte **T+1**, mientras que para **T+2** los resultados son más satisfactorios al emplear **un solo factor**.
- Al **excluir el periodo asociado a la COVID-19**, se observa que los errores (**en variaciones mensuales**) disminuyen a **0.52** y **0.53** para **T+1** y **T+2** respectivamente y que el **uso de un factor** continúa ofreciendo los **mejores resultados** en términos de precisión.
- Finalmente, al emplear **ventanas expansivas**, los errores de estimación **aumentan ligeramente**, aunque las **conclusiones generales se mantienen estables**.



CONCLUSIONES

- Los indicadores IOAE e IOCP generan **estimaciones oportunas y confiables**, en línea con las experiencias metodológicas reportadas por la **academia**, los **centros de investigación** y diversas **oficinas nacionales de estadística**.
- El IOAE, y según una interpretación del análisis de sentimientos, se ha consolidado como una herramienta utilizada por los usuarios con **finés analíticos y de toma de decisiones objetivas**.
- Asimismo, la **metodología econométrica/estadística** demuestra una **robustez significativa** frente a otras técnicas de *nowcasting*, particularmente para el horizonte **T+2**.
- Los resultados de esta investigación fueron publicados en ***Latin American Economic Review***, <https://www.latinaer.org/index.php/laer/article/view/461>.

RECOMENDACIONES

- En consecuencia, se **recomienda la adopción y uso de estas metodologías** como herramientas de apoyo en la **toma de decisiones económicas en México, siempre y cuando estén validadas estadísticamente.**
- También, es valioso realizar un **análisis de sentimientos** para el resto de **indicadores oportunos** del INEGI y **evaluaciones econométricas** en tiempo pseudo-real para **verificar los resultados** en ventanas de tiempo más grandes y respecto a otras técnicas.
- Finalmente, se considera **prioritario mantener el monitoreo continuo del error de estimación** a lo largo del tiempo, a fin de asegurar la estabilidad y confiabilidad del sistema de indicadores oportunos.



¡GRACIAS!

CONOCIENDO
**MÉ
XI
CO**

800 111 46 34

www.inegi.org.mx

atencion.usuarios@inegi.org.mx



INEGIINFORMA

