



MEDIO AMBIENTE
SECRETARÍA DE MEDIO AMBIENTE Y RECURSOS NATURALES



IMTA
INSTITUTO MEXICANO
DE TECNOLOGÍA DEL AGUA

**COORDINACIÓN DE GOBERNANZA DEL AGUA Y
FORTALECIMIENTO DE CAPACIDADES**

SUBCOORDINACIÓN DE POSGRADO Y EDUCACIÓN CONTINUA

T E S I S

**IMPLEMENTACIÓN DE UNA METODOLOGÍA DE CORRECCIÓN DE
SESGO PARA PRONÓSTICO METEOROLÓGICO NUMÉRICO**

que para obtener el grado de
**Maestro en Ciencias y Tecnología del Agua:
Hidrometeorología**

presenta
Flavio César Acosta Barradas

Tutor: M. C. Roberto Ramírez Villa
Cotutora: Dra. Mirce Ivón Morales Velázquez

Jiutepec, Morelos

Año: 2023

F . C 4 . 0 3 . 4 4 . 1

Paseo Cuauhnáhuac No. 8532, Col. Progreso, C.P. 62550, Jiutepec, Morelos.
Tel.: (777) 329 36 00 www.gob.mx/imta





Con fundamento en los artículos 21 y 27 de la Ley Federal del Derecho de Autor y como titular de los derechos moral y patrimoniales de la obra titulada “IMPLEMENTACIÓN DE UNA METODOLOGÍA DE CORRECCIÓN DE SESGO PARA PRONÓSTICO METEOROLÓGICO NUMÉRICO”, otorgo de manera gratuita y permanente al Instituto Mexicano de Tecnología del Agua, autorización para que fijen la obra en cualquier medio, incluido el electrónico, y la divulguen entre su personal, estudiantes o terceras personas, sin que pueda percibir por tal divulgación una contraprestación.

Flavio César Acosta Barradas

Jiutepec, Morelos, junio del 2023

F.C4.03.43.1



COMITÉ TUTORIAL:

Dr. Julio Sergio Santana Sepúlveda
Dr. René Lobato Sánchez
M. C. Roberto Ramírez Villa
Dr. Víctor Kevin Contreras Tereza
Dra. Mirce Ivón Morales Velázquez

Director de tesis:
M.C. Roberto Ramírez Villa



Co-directora de tesis:
Dra. Mirce Ivón Morales Velázquez



AGRADECIMIENTOS Y DEDICATORÍA

A mi familia: a mi madre, María Yolanda Barradas Herrera; a mi padre, Julio César Acosta López; a mi Hermano, Julio César Acosta Barradas; y a mi novia, Arisbeth Cobos Loyo, por su apoyo en cada una de las etapas de mi vida.

Extiendo mi sincero agradecimiento a mis tutores y miembros del comité tutorial: el M.C. Roberto Ramírez Villa, la Dra. Mirce Ivón Morales Velázquez, el Dr. Víctor Kevin Contreras Tereza, el Dr. René Lobato Sánchez y el Dr. Julio Sergio Santana Sepúlveda. Agradezco sus valiosas orientaciones, sugerencias y recomendaciones que, durante estos dos años, resultaron fundamentales para elevar la calidad de la investigación.

Quiero expresar mi agradecimiento a mi compañero y amigo, Fabián Salazar Vázquez, así como a todos los compañeros que tuve el privilegio de conocer a lo largo de estos cuatro semestres.

Agradezco sinceramente a todas las personas que brindaron su apoyo en las consultas realizadas, contribuyendo significativamente al avance y éxito de la investigación.

Por último, agradezco al Instituto Mexicano de Tecnología del Agua por brindarme la oportunidad de cursar el posgrado. Asimismo, agradezco a las personas que forman parte de esta institución y contribuyen con su labor.

ÍNDICE

Capítulo 1. Introducción.....	1
Antecedentes.....	2
1.1.1 Pronóstico meteorológico y estado del tiempo	2
1.1.2 Modelación numérica de la atmósfera	3
1.1.3 Asimilación de datos	7
1.1.4 Técnicas de postprocesamiento estadístico	13
Planteamiento del problema	18
Objetivos.....	19
1.2.1 Objetivo general.....	19
1.2.2 Objetivos específicos	19
Descripción de la región de estudio.....	20
Justificación	25
Alcances y limitaciones	26
Hipótesis	28
Equipo de cómputo y lenguaje de programación	28
Capítulo 2. Metodología.....	30
Métricas de evaluación.....	30
Método de corrección basado en filtros de Kalman	32
Capítulo 3. Desarrollo	36
Datos meteorológicos utilizados: observaciones y pronósticos	36
2.3.1 Datos observados de estaciones climáticas convencionales	36
2.3.2 Datos pronosticados por el sistema SREF	37
Aplicación de métodos estadísticos para identificación de sesgo	39
Aplicación del método de corrección de sesgo	52
Capítulo 4. Análisis de resultados	62
Identificación de reducción de sesgo al aplicar Filtro de Kalman Adaptativo.....	81
Capítulo 5. Conclusiones y recomendaciones	92
Conclusiones.....	92
Recomendaciones	96
Referencias	99
ANEXOS	104
Anexo 1: Acrónimos.....	105
Anexo 2: RMSE, MAE y COE, pronóstico vs observación.....	107

Anexo 3: RMSE, MAE y COE, corrección vs observación.....	112
Análisis de los resultados mostrados en el Anexo 2 y 3.....	124

ÍNDICE DE TABLAS

Tabla 1: Ejemplo hipotético de la manera de realizar el cálculo de la temperatura máxima y mínima el pronóstico del sistema SREF.....	38
Tabla 2: Intervalos de sesgo para la temperatura máxima en el trimestre de diciembre, enero y febrero (DEF) 2020-2021, pronóstico vs observación.....	41
Tabla 3: Intervalos de sesgo para la temperatura máxima en el trimestre de mayo, junio y julio 2020, pronóstico vs observación.....	42
Tabla 4: Intervalos de sesgo para la temperatura mínima en el trimestre de diciembre, enero y febrero 2020-2021, pronóstico vs observación.....	43
Tabla 5: Intervalos de sesgo para la temperatura mínima en el trimestre de mayo, junio y julio 2020, pronóstico vs observación.....	44
Tabla 6: Intervalos de sesgo para la precipitación en el trimestre de diciembre, enero y febrero 2020-2021, pronóstico vs observación.....	45
Tabla 7: Intervalos de sesgo para la precipitación en el trimestre de mayo, junio y julio 2020, pronóstico vs observación.....	46
Tabla 8: Valor de sesgo máximo positivo y negativo encontrado para cada caso de estudio, pronóstico vs observación.....	49
Tabla 9: Resumen de sesgos e intervalos para las variables de interés.....	49
Tabla 10 : Distribución del sesgo positivo y negativo para cada una de las variables y periodos analizados, identificación del comportamiento del sistema SREF.....	50
Tabla 11: Intervalos de sesgo para la temperatura máxima en el trimestre de diciembre, enero y febrero 2020-2021, pronóstico vs observación.....	82
Tabla 12: Intervalos de sesgo para la temperatura máxima en el trimestre de mayo, junio y julio 2020, pronóstico vs observación.....	83
Tabla 13: Intervalos de sesgo para la temperatura mínima en el trimestre de diciembre, enero y febrero 2020-2021, pronóstico vs observación.....	84
Tabla 14: Intervalos de sesgo para la temperatura mínima en el trimestre de mayo, junio y julio 2020, corrección vs observación.....	85
Tabla 15: Intervalos de sesgo para la precipitación en el trimestre de diciembre, enero y febrero 2020-2021, corrección vs observación.....	86
Tabla 16: Intervalos de sesgo para la precipitación en el trimestre de mayo, junio y julio 2020, corrección vs observación.....	87
Tabla 17: Valor de máximo sesgo negativo y positivo encontrado para cada caso de estudio, corrección vs observación.....	89
Tabla 18: Resumen de sesgos e intervalos para las variables de interés, corrección vs observación.....	89
Tabla 19: Distribución del sesgo positivo y negativo para cada una de las variables y periodos analizados, identificación del comportamiento de la corrección por parte del Filtro de Kalman Adaptativo.....	89
Tabla 20: Intervalos de RMSE para la temperatura máxima en el trimestre de diciembre, enero y febrero 2020-2021, pronóstico vs observación.....	109

Tabla 21: Intervalos de RMSE para la temperatura máxima en el trimestre de mayo, junio y julio 2020, pronóstico vs observación	109
Tabla 22: . Intervalos de RMSE para la temperatura mínima en el trimestre de diciembre, enero y febrero 2020-2021, pronóstico vs observación.....	110
Tabla 23: Intervalos de RMSE para la temperatura mínima en el trimestre de mayo, junio y julio 2020, pronóstico vs observación.	110
Tabla 24: Intervalos de MAE para la temperatura máxima en el trimestre de diciembre, enero y febrero 2020-2021, pronóstico vs observación.....	110
Tabla 25: Intervalos de MAE para la temperatura máxima en el trimestre de mayo, junio y julio 2020, pronóstico vs observación.	111
Tabla 26: Intervalos de MAE para la temperatura mínima en el trimestre de diciembre, enero y febrero 2020-2021, pronóstico vs observación.....	111
Tabla 27: Intervalos de MAE para la temperatura mínima en el trimestre de mayo, junio y julio 2020, pronóstico vs observación.	111
Tabla 28: Intervalos de RMSE para la temperatura máxima en el trimestre de diciembre, enero y febrero 2020-2021, corrección vs observación.....	113
Tabla 29: Intervalos de RMSE para la temperatura máxima en el trimestre de mayo, junio y julio 2020, corrección vs observación.	114
Tabla 30: Intervalos de RMSE para la temperatura mínima en el trimestre de diciembre, enero y febrero 2020-2021, corrección vs observación.....	114
Tabla 31: Intervalos de RMSE para la temperatura mínima en el trimestre de mayo, junio y julio 2020, corrección vs observación.	114
Tabla 32: Intervalos de MAE para la temperatura máxima en el trimestre de diciembre, enero y febrero 2020-2021, corrección vs observación.....	114
Tabla 33: Intervalos de MAE para la temperatura máxima en el trimestre de mayo, junio y julio 2020, corrección vs observación.	114
Tabla 34: Intervalos de MAE para la temperatura mínima en el trimestre de diciembre, enero y febrero 2020-2021, corrección vs observación.....	115
Tabla 35: Intervalos de MAE para la temperatura mínima en el trimestre de mayo, junio y julio 2020, corrección vs observación.	115
Tabla 36: Coeficiente de correlación de Pearson promedio encontrado para cada caso de estudio.....	116

ÍNDICE DE FIGURAS

Figura 1: Dominios del sistema Short-Range Ensemble Forecast (SREF). (COMET-MetEd, 2015).....	4
Figura 2: Distribución de estaciones climáticas convencionales, a lo largo del territorio mexicano. Elaboración propia mediante QGIS.	20
Figura 3: Temperatura media anual en el territorio mexicano. Obtenido de Coll-Hurtado, 1990.	22
Figura 4: Histograma de sesgo promedio para la temperatura máxima en el periodo DEF 2020-2021, pronóstico vs observación.	41
Figura 5: Histograma de sesgo promedio para la temperatura máxima en el periodo MJJ 2020, pronóstico vs observación.	42
Figura 6: Histograma de sesgo promedio para la temperatura mínima en el periodo DEF	

2020-2021, pronóstico vs observación.	43
Figura 7: Histograma de sesgo promedio para la temperatura mínima en el periodo MJJ 2020, pronóstico vs observación.	44
Figura 8: Histograma de sesgo promedio para la precipitación en el periodo DEF 2020-2021, pronóstico vs observación.	45
Figura 9: Histograma de sesgo promedio para la precipitación en el periodo MJJ 2020, pronóstico vs observación.	47
Figura 10: Ecuaciones del Filtro de Kalman Adaptativo utilizadas para realizar la corrección de sesgo.	53
Figura 11: Algoritmo de Filtro de Kalman Adaptativo escrito en Python.	54
Figura 12: Ejemplo de los resultados obtenidos al aplicar el código en Python.	55
Figura 13: Esquemización de la manera en la que se calculó la covarianza del error de la ecuación de estado $Q_{t-k\Delta t}$ y la covarianza del error de la ecuación de observación R_t , según lo propuesto por (Galanis & Anadranistokis , 2002).	56
Figura 14: Sesgo de la temperatura máxima entre el pronóstico del SREF y las observaciones de 1570 estaciones climáticas convencionales durante diciembre, enero y febrero de 2020-2021. Representación de sesgos por colores: rojo para sesgos positivos, blanco para para sesgos cercanos a cero y azul para sesgos negativos. Elaboración propia mediante el uso de QGIS.	63
Figura 15: Sesgo de la temperatura máxima entre la corrección mediante Filtro de Kalman Adaptativo y las observaciones de 1570 estaciones climáticas convencionales durante diciembre, enero y febrero de 2020-2021. Representación de sesgos por colores: rojo para sesgos positivos, blanco para sesgos cercanos a cero y azul para sesgos negativos. Elaboración propia mediante el uso de QGIS.	64
Figura 16: Sesgo de la temperatura máxima entre el pronóstico del SREF y las observaciones de 1570 estaciones climáticas convencionales durante mayo, junio y julio de 2020. Representación de sesgos por colores: rojo para sesgos positivos, blanco para sesgos cercanos a cero y azul para sesgos negativos. Elaboración propia mediante el uso de QGIS.	65
Figura 17: Sesgo de la temperatura máxima entre la corrección mediante Filtro de Kalman Adaptativo y las observaciones de 1570 estaciones climáticas convencionales durante mayo, junio y julio de 2020. Representación de sesgos por colores: rojo para sesgos positivos, blanco para sesgos cercanos a cero y azul para sesgos negativos. Elaboración propia mediante el uso de QGIS.	66
Figura 18: Sesgo de la temperatura mínima entre el pronóstico del SREF y las observaciones de 1570 estaciones climáticas convencionales durante diciembre, enero y febrero de 2020-2021. Representación de sesgos por colores: rojo para sesgos positivos, blanco para sesgos cercanos a cero y azul para sesgos negativos. Elaboración propia mediante el uso de QGIS.	67
Figura 19: Sesgo de la temperatura mínima entre la corrección mediante Filtro de Kalman Adaptativo y las observaciones de 1570 estaciones climáticas convencionales durante diciembre, enero y febrero de 2020-2021. Representación de sesgos por colores: rojo para sesgos positivos, blanco para sesgos cercanos a cero y azul para sesgos negativos. Elaboración propia mediante el uso de QGIS.	68
Figura 20: Sesgo de la temperatura mínima entre el pronóstico del SREF y las observaciones de 1570 estaciones climáticas convencionales durante mayo, junio y julio de 2020. Representación de sesgos por colores: rojo para sesgos positivos, blanco para sesgos	68

cercanos a cero y azul para sesgos negativos. Elaboración propia mediante el uso de QGIS.	69
Figura 21: Sesgo de la temperatura mínima entre la corrección mediante Filtro de Kalman Adaptativo y las observaciones de 1570 estaciones climáticas convencionales durante mayo, junio y julio de 2020. Representación de sesgos por colores: rojo para sesgos positivos, blanco para sesgos cercanos a cero y azul para sesgos negativos. Elaboración propia mediante el uso de QGIS.....	70
Figura 22: Sesgo de la precipitación entre el pronóstico del SREF y las observaciones de 1570 estaciones climáticas convencionales durante diciembre, enero y febrero de 2020-2021. Representación de sesgos por colores: verdes para sesgos positivos, blanco para sesgos cercanos a cero y morado para sesgos negativos. Elaboración propia mediante el uso de QGIS.....	71
Figura 23: Sesgo de precipitación entre la corrección mediante Filtro de Kalman Adaptativo y las observaciones de 1570 estaciones climáticas convencionales durante diciembre, enero y febrero de 2020-2021. Representación de sesgos por colores: verde para sesgos positivos, blanco para sesgos cercanos a cero y morado para sesgos negativos. Elaboración propia mediante el uso de QGIS.....	72
Figura 24: Sesgo de la precipitación entre el pronóstico del SREF y las observaciones de 1570 estaciones climáticas convencionales durante mayo, junio y julio de 2020. Representación de sesgos por colores: verde para sesgos positivos, blanco para sesgos cercanos a cero y morado para sesgos negativos. Elaboración propia mediante el uso de QGIS.....	73
Figura 25: Sesgo de la precipitación entre la corrección mediante Filtro de Kalman Adaptativo y las observaciones de 1570 estaciones climáticas convencionales durante mayo, junio y julio de 2020. Representación de sesgos por colores: verdes para sesgos positivos, blanco para sesgos cercanos a cero y morado para sesgos negativos. Elaboración propia mediante el uso de QGIS.....	74
Figura 26: Comparación entre datos de temperatura máxima extraídos de la media del ensamble SREF (azul) para las coordenadas de la estación de Jiutepec, Mor., temperaturas máximas corregidas (naranja) con Filtro de Kalman Adaptado y observaciones (verde), periodo DEF.	75
Figura 27: Comparación entre datos de temperatura máxima extraídos de la media del ensamble SREF (azul) para las coordenadas de la estación de Jiutepec, Mor., temperaturas máximas corregidas (naranja) con Filtro de Kalman Adaptado y observaciones (verde), periodo MJJ.	76
Figura 28: Comparación entre datos de temperatura mínima extraídos de la media del ensamble SREF (azul) para las coordenadas de la estación de Jiutepec, Mor., temperaturas máximas corregidas (naranja) con Filtro de Kalman Adaptado y observaciones (verde), periodo DEF.	77
Figura 29: Comparación entre datos de temperatura mínima extraídos de la media del ensamble SREF (azul) para las coordenadas de la estación de Jiutepec, Mor., temperaturas máximas corregidas (naranja) con Filtro de Kalman Adaptado y observaciones (verde), periodo MJJ.	78
Figura 30: Comparación entre datos de precipitación extraídos de la media del ensamble SREF (azul) para las coordenadas de la estación de Jiutepec, Mor., precipitaciones corregidas (naranja) con Filtro de Kalman Adaptado y observaciones (verde), periodo DEF.	79

Figura 31: Comparación entre datos de precipitación extraídos de la media del ensamble SREF (azul) para las coordenadas de la estación de Jiutepec, Mor., precipitaciones corregidas (naranja) con Filtro de Kalman Adaptado y observaciones (verde), periodo MJJ.	80
Figura 32: Histograma de sesgo promedio para la temperatura máxima en el periodo DEF 2020-2021, corrección vs observación.	82
Figura 33: Histograma de sesgo promedio para la temperatura máxima en el periodo MJJ 2020, corrección vs observación.	83
Figura 34: Histograma de sesgo promedio para la temperatura mínima en el periodo DEF 2020-2021, corrección vs observación.	84
Figura 35: Histograma de sesgo promedio para la temperatura mínima en el periodo MJJ 2020, corrección vs observación.	85
Figura 36: Histograma de sesgo promedio para la Precipitación en el periodo DEF 2020-2021, corrección vs observación.	86
Figura 37: Histograma de sesgo promedio para la Precipitación en el periodo MJJ 2020, corrección vs observación.	87
Figura 38: Comparación de la dispersión entre la temperatura máxima corregida (izquierda) y la pronosticada (derecha), contra las observaciones para la estación de Jiutepec, Mor., periodo DEF.	118
Figura 39: Comparación de la dispersión entre la temperatura máxima corregida (izquierda) y la pronosticada (derecha), contra las observaciones para la estación de Jiutepec, Mor., periodo MJJ.	119
Figura 40: Comparación de la dispersión entre la temperatura mínima corregida (izquierda) y la pronosticada (derecha), contra las observaciones para la estación de Jiutepec, Mor., periodo DEF.	120
Figura 41: Comparación de la dispersión entre la temperatura mínima corregida (izquierda) y la pronosticada (derecha), contra las observaciones para la estación de Jiutepec, Mor., periodo MJJ.	121
Figura 42: Comparación de la dispersión entre la precipitación corregida (izquierda) y la pronosticada (derecha), contra las observaciones para la estación de Jiutepec, Mor., periodo DEF.	122
Figura 43: Comparación de la dispersión entre la precipitación corregida (izquierda) y la pronosticada (derecha), contra las observaciones para la estación de Jiutepec, Mor., periodo MJJ.	123

Resumen

Este estudio se enfoca en la corrección del sesgo en los pronósticos meteorológicos numéricos a través del uso de una técnica Model Output Statistics (MOS) adaptativa basada en filtros de Kalman. Se aplica esta técnica como postprocesamiento de los resultados numéricos de temperatura máxima (°C), temperatura mínima (°C) y precipitación (mm), de la media del ensamble del sistema Short-Range Ensemble Forecast (SREF). El objetivo de esta tesis es implementar una metodología de corrección de sesgo para reducir las diferencias entre los pronósticos del sistema SREF y observaciones puntuales. Se utilizan observaciones de 1570 estaciones climáticas convencionales distribuidas en el territorio mexicano durante los trimestres de (diciembre, enero y febrero) y (mayo, junio y julio) de los años 2020-2021. Se aplican diversas métricas estadísticas como el BIAS, RMSE, MAE y COE, para identificar el sesgo del sistema SREF y para evaluar la efectividad de la corrección mediante la aplicación del Filtro de Kalman Adaptativo. De acuerdo con los resultados obtenidos en este trabajo, se demostró que la metodología propuesta es factible y útil en la corrección de sesgo del sistema SREF en las variables de temperatura máxima y temperatura mínima en el territorio mexicano, debido a que se comprobó que la metodología logró reducir el sesgo con valores muy cercanos a cero. En cuanto a la variable de precipitación, los resultados fueron favorables, pero la disminución no fue tan significativa como para el caso de la temperatura. Los resultados obtenidos permiten comprender el impacto de las técnicas de postprocesamiento en la mejora de los pronósticos numéricos del SREF al reducir las diferencias entre los valores estimados por el sistema y las observaciones.

Palabras Clave: Filtro de Kalman Adaptativo, SREF, corrección de sesgo, postprocesamiento, modelos de simulación numérica del tiempo.

Capítulo 1. Introducción

En el ámbito de la meteorología, el pronóstico del tiempo desempeña un papel fundamental en la toma de decisiones en diversos sectores. La modelación numérica de la atmósfera, a través de los modelos de predicción numérica del tiempo (NWP, por su acrónimo en inglés), se ha convertido en una herramienta esencial para la generación de pronósticos. Sin embargo, estos modelos presentan errores sistemáticos y no sistemáticos que afectan su precisión. Estos errores pueden deberse a diversas razones, una de ellas es la limitación de la capacidad del modelo para simular las propiedades físicas y dinámicas del sistema a su escala de resolución. Además, al comparar los resultados de los modelos NWP con observaciones locales, se han encontrado diferencias significativas.

Ante este desafío, las técnicas estadísticas de postprocesamiento han demostrado ser útiles para reducir los errores presentes en los modelos NWP. En este trabajo, se propone la implementación de una técnica MOS adaptativa basada en filtros de Kalman, denominada Filtro de Kalman Adaptativo, que permite postprocesar secuencialmente los pronósticos, actualizando continuamente los parámetros de corrección a medida que se obtienen nuevas observaciones. La técnica MOS, es un tipo de postprocesamiento estadístico que se utiliza para mejorar la capacidad de pronóstico de los modelos meteorológicos numéricos al relacionar los resultados del modelo con observaciones. De acuerdo con Glahn y Lowry (1972), MOS es una técnica objetiva de pronóstico del tiempo que consiste en determinar una relación estadística entre un predictando y las variables pronosticadas por un modelo numérico en algún tiempo de proyección.

El objetivo principal de esta investigación es corregir el sesgo de las salidas del sistema SREF. Se han considerado las variables de temperatura máxima, temperatura mínima y precipitación. Para ello, se ha realizado una comparación entre las correcciones aplicadas por el filtro y las observaciones de 1570 estaciones climáticas convencionales distribuidas en el territorio mexicano. Se analizaron los resultados durante los trimestres de MJJ e DEF de los años 2020-2021. Mediante el cálculo de métricas estadísticas, se busca evaluar el impacto de la corrección realizada por el Filtro de Kalman Adaptativo y determinar si se logra una reducción

significativa del sesgo en las series corregidas. Los resultados obtenidos proporcionarán información valiosa sobre la mejora de la precisión de los pronósticos del sistema SREF a través de este enfoque de postprocesamiento.

Antecedentes

1.1.1 Pronóstico meteorológico y estado del tiempo

Se le conoce como estado del tiempo, al conjunto de meteoros o condiciones atmosféricas presentes en un lugar o momento determinado (Garduño, 1994). El estado del tiempo no es constante, varía a lo largo del año y puede ser caracterizado a partir de la medición de variables meteorológicas tales como temperatura, velocidad y dirección del viento, precipitación, presión barométrica, humedad, entre otras. Así como, con la presencia de meteoros o fenómenos meteorológicos, por ejemplo, granizo, nieve, niebla, tormentas, sistemas frontales, suradas, ciclones tropicales y demás. Estas variaciones en el estado del tiempo a lo largo del año son moduladas por distintos factores relacionados con los mecanismos físicos que mantienen el balance de energía del planeta y el cambio de las estaciones ocasionado por la inclinación del eje terrestre y la manera en la que incide la radiación solar en la superficie, entre otros.

El estado del tiempo para un estado futuro puede ser estimado a partir del análisis de diversos datos, a esto se le conoce como pronóstico meteorológico. Una forma de predecir el estado futuro de la atmósfera es utilizando aproximaciones numéricas de las ecuaciones de la dinámica de fluidos geofísicos. Combinando de forma óptima estas estimaciones numéricas de corto plazo con observaciones se puede hacer un buen pronóstico meteorológico (Dillon et al., s. f.), sumado a la habilidad del meteorólogo operativo para realizar la interpretación de la información disponible. Cabe mencionar, que cuanto mejor sea el diagnóstico de las condiciones actuales de la atmósfera, mejor será la estimación de las condiciones para inicializar los modelos numéricos del tiempo y por lo tanto mejor será la calidad del pronóstico meteorológico numérico (Warner et al., 1989; Harms et al., 1992; Kalnay, 2003).

Contar con pronósticos precisos del estado del tiempo es de vital importancia para la sociedad,

ya que sirven como base fundamental para la toma de decisiones en diversos ámbitos como la salud, agricultura, transporte, pesca, actividades artísticas y el académico. La calidad de los pronósticos meteorológicos influye directamente en la planificación de actividades, la gestión de recursos y la seguridad de las personas. Por lo tanto, es esencial contar con pronósticos confiables y para garantizar un desarrollo sostenible en todas estas áreas.

1.1.2 Modelación numérica de la atmósfera

En la actualidad, los modelos de simulación numérica del tiempo son una herramienta esencial para la generación de pronósticos meteorológicos. En Meteorología, el objetivo de los llamados modelos de predicción numérica consiste en calcular matemáticamente la evolución temporal del flujo atmosférico a partir del conocimiento de sus características en un determinado momento. Esto se realiza mediante la resolución de las ecuaciones de continuidad, movimiento, energía termodinámica y humedad, conocidas bajo el nombre de ecuaciones primitivas, que son expresiones basadas respectivamente en el principio de conservación de masa, cantidad de movimiento, calor y vapor de agua en un volumen elemental de aire. (Castro M., Gonzales R., Portela A., 1991).

Los modelos numéricos, producen hoy en día pronósticos tridimensionales de un estado futuro de la atmósfera mediante una estimación del presente de esta. La precisión del pronóstico producido radica fundamentalmente en una buena descripción del estado inicial de la atmósfera y del grado de conocimiento de la física que describe los procesos atmosféricos. A este estado inicial se le conoce como análisis, y es creado por una combinación óptima de datos observados y pronósticos de corto plazo de un análisis previo, a través de un proceso que se le conoce como asimilación de datos. Por lo tanto, para producir un análisis y consecuentemente un buen pronóstico es necesario disponer de gran cantidad de observaciones de gran calidad. (Lighezzolo R. A., 2014).

En el mundo, hay una amplia variedad de modelos de predicción numérica de la atmósfera, y uno de ellos es el Weather Research and Forecasting model (WRF). Ha sido desarrollado con el propósito de realizar investigaciones y generar pronósticos operativos del tiempo. Fue

diseñado principalmente por el National Center for Atmospheric Research (NCAR), la National Oceanic and Atmospheric Administration (NOAA), el National Center for Environmental Prediction (NCEP), el Earth System Research Laboratory (ESRL), la Air Force Weather Agency (AFWA), el Naval Research Laboratory (NLR) el Center for Analysis and Prediction Storms (CAPS) y la Federal Aviation Administration (FAA), actualmente se ejecuta en el NCEP (Grupo de Interacción Océano Atmósfera, 2017). El modelo WRF cuenta con dos núcleos principales (Advanced Research WPS (ARW) y Nonhydrostatical Mesoscale Model (NMM), un sistema de asimilación de datos y una arquitectura de software que permite la paralelización de procesos. El modelo permite un amplio uso de aplicaciones meteorológicas en escalas con rangos que van desde unos pocos kilómetros hasta miles de kilómetros. El desarrollo del WRF comenzó a principios de los años 90's. (Lighezzolo R. A., 2014).

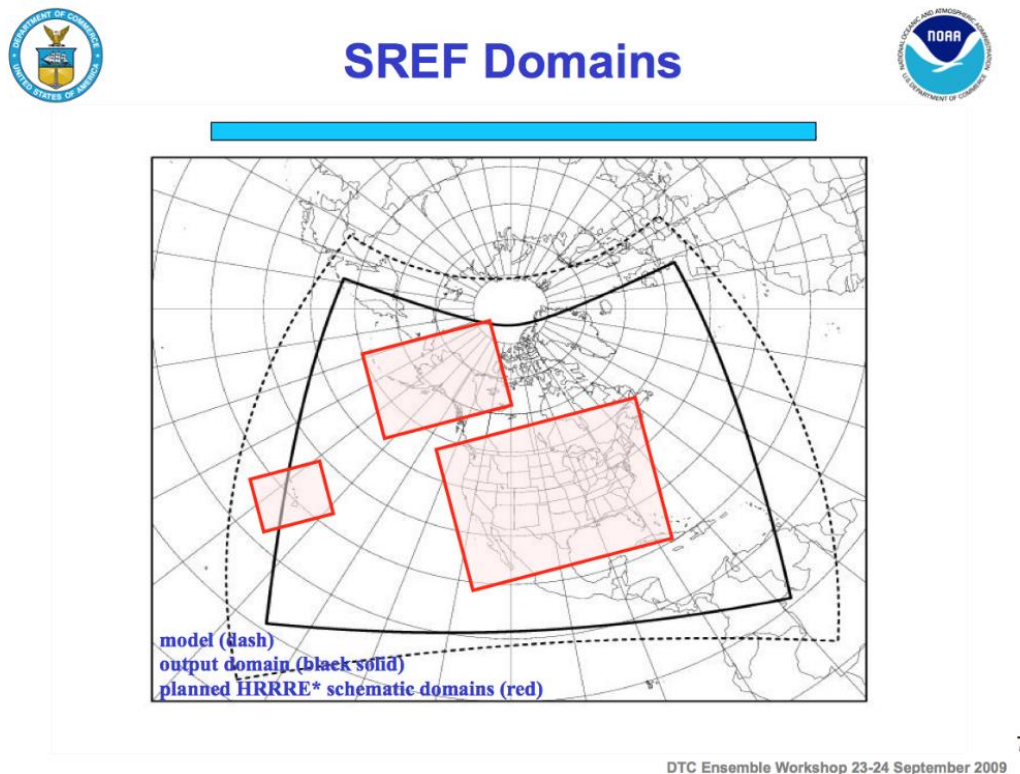


Figura 1: Dominios del sistema Short-Range Ensemble Forecast (SREF). (COMET-MetEd, 2015).

Para el desarrollo de este trabajo, se utilizarán salidas del sistema Short-Range Ensemble

Forecast (SREF). Debido a que NCEP no almacena históricos por el gran volumen de información generada, se han almacenado y respaldado estas salidas. Para incorporar un enfoque estocástico a un modelo determinístico como el WRF, se utilizarán ensambles. En nuestro caso, se trabajará utilizando la media del ensamble, siguiendo las pautas proporcionadas por el Centro Europeo para Pronóstico Meteorológicos a Mediano Plazo (ECMWF, 2015), que sugiere aplicar el Filtro de Kalman Adaptativo al pronóstico de control o a la media del ensamble. Esto se debe a que, si se aplicara la técnica a cada miembro del ensamble, las anomalías de pronóstico representadas por la dispersión del conjunto se tratarían erróneamente como errores del modelo, lo que resultaría en una atenuación de la dispersión del conjunto después del postprocesamiento (Pelosi, 2017).

SREF es un sistema probabilístico que utiliza dos núcleos del modelo WRF, tales son: Advanced Research WRF (WRF-ARW) y NEMS Non-hydrostatic Multiscale Model on the B grid (NMMB), de los cuales, para cada uno de los modelos se realiza una ejecución de control y un conjunto de 12 miembros perturbados, lo que resulta en un ensamble de 26 miembros perturbando las condiciones iniciales (COMET-MetEd, 2015a). Estas condiciones iniciales son proporcionadas por tres modelos determinísticos: The North American Meso-scale Model (NAM) Data Assimilation System (NDAS), Global Forecast System (GFS) Data Assimilation System (GDAS) y Rapid Refresh (RAP) (COMET-MetEd, 2015b), además de perturbar condiciones iniciales también se modifican parametrizaciones físicas como convección, capa límite planetaria, radiación y microfísica. Asimismo, en lo que respecta a la resolución vertical, el modelo WRF-ARW utiliza una coordenada sigma que sigue el terreno, mientras que el NMMB utiliza una coordenada vertical híbrida sigma-presión que sigue el terreno desde la superficie hasta la cima del modelo a 50-hPa (COMET-MetEd, 2015c).

En cuanto a las características de las salidas del modelo de ensamble WRF-SREF se considera el dominio para Norteamérica (aunque el ensamble también maneja los dominios CONUS (parte continental de Estados Unidos), Alaska y Hawai) (COMET-MetEd, 2015d), tiene un espaciamiento en tiempo de 3 horas, ejecución diaria (03, 09, 15, 21 UTC), resolución espacial de 16 km entre cada punto de malla (COMET-MetEd, 2015e) y un alcance de pronóstico

aproximadamente de tres días. Si desea conocer más detalladamente las características del sistema SREF, se sugiere consultar la fuente oficial: <https://sites.google.com/ucar.edu/operational-models-encyclo/probabilistic-models/sref>, en ella se describen ampliamente las particularidades mencionadas sobre SREF en este trabajo.

Los datos derivados de la salida de SREF¹ están disponibles en formato grib2 con salidas disponibles cada 6 horas, para acceder a ellos es mediante FTP y HTTPS. Para esta investigación, la descarga y procesamiento de datos se llevó a cabo en una estación de trabajo llamada Aurora, cuyas características computacionales se mencionan más adelante.

En dicha estación se comenzaron a respaldar datos a partir del 01 de abril del 2020 de las variables: precipitación total en superficie (kg/m^2), energía potencial convectiva disponible CAPE (J/kg), inhibición convectiva en superficie (J/kg), temperatura de punto de rocío (2 metros sobre el suelo, 300mb, 500mb, 700mb y 850mb) (K), presión reducida al nivel del mar (Pa), humedad relativa (2 metros sobre el suelo, 300mb, 500mb, 700mb y 850mb) (%), temperatura a 2 metros del suelo (K) y componente U y V del viento (10 metros sobre el suelo, 250mb, 300mb, 500mb, 600mb, 700mb, 850mb, 1000mb) (m/s).

Es fundamental destacar que los modelos numéricos del tiempo contienen errores tanto sistemáticos como aleatorios al pronosticar variables meteorológicas. Uno de los motivos detrás de este problema es la ineficiente estimación del estado actual del flujo atmosférico. Para abordar esta situación, se han desarrollado métodos que permiten estimar con mayor precisión el estado del flujo atmosférico, conocidos como “asimilación de datos”. Adicionalmente, se utilizan técnicas estadísticas de postprocesamiento para reducir el error sistemático de los modelos numéricos. En las siguientes secciones, se describirán ambos enfoques con más detalle.

¹ <https://www.nco.ncep.noaa.gov/pmb/products/sref/>

1.1.3 Asimilación de datos

Talagrand (1997) define “asimilación de datos” como el uso de toda la información disponible para conocer con mayor precisión el estado del flujo atmosférico u oceánico. Este estado atmosférico u oceánico se puede estimar mediante la combinación óptima de pronósticos de corto plazo (proveniente de modelos numéricos) y observaciones. Warner et al. (1989); Harms et al (1992); Kalnay (2003) señalan que la calidad del pronóstico mejora a medida que se mejora la estimación de las condiciones iniciales.

En la actualidad, el estado del arte de los métodos de asimilación de datos comprende los métodos variacionales y los métodos secuenciales. Los métodos variacionales utilizan algoritmos de optimización para minimizar una función objetivo predefinida que mide la diferencia entre el modelo y las observaciones, como el 3DVAR (Three-dimensional variational assimilation) y 4DVAR (Four-dimensional variational assimilation). Los métodos secuenciales, como el Filtro de Kalman (KF) en sus diferentes versiones (Beck et al., 2016), proporcionan estimaciones del estado del sistema de manera secuencial, propagando la información hacía adelante en el tiempo. Estos métodos aprovechan las observaciones a medida que van apareciendo para mejorar la capacidad de predicción, y permiten que el modelo se actualice continuamente. Además, proporcionan un marco probabilístico necesario para cuantificar la incertidumbre asociada a los resultados. De esta manera, el estado del sistema se va actualizando y mejorando en forma secuencial, y la información se va propagando en el tiempo. Ambos métodos se describen a continuación.

Métodos Variacionales

En 1958, Sasaki, desarrolló un método de asimilación de datos basado en el cálculo variacional (3DVAR). Este método trata de encontrar el valor de \mathbf{x} que minimiza la función de coste u objetivo $J(\mathbf{x})$. A continuación, se muestra la función objetivo base a minimizar, de los métodos variacionales:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b) + \frac{1}{2}(\mathbf{H}(\mathbf{x}) - \mathbf{y})^T \mathbf{R}^{-1}(\mathbf{H}(\mathbf{x}) - \mathbf{y}) \dots (1)$$

Donde \mathbf{B} y \mathbf{R} son errores, \mathbf{R} equivale al error de las observaciones y \mathbf{B} al error del Background (primera estimación) y \mathbf{x}^b serían las condiciones iniciales de la primera estimación, \mathbf{y} son observaciones y H es un operador de observación.

Al minimizar la ecuación objetivo, se obtiene una estimación óptima del estado del sistema que combina la información del modelo numérico del tiempo y las observaciones. La minimización numérica, aunque puede resultar complicada, se puede lograr utilizando técnicas de optimización, por ejemplo: métodos gradiente decreciente, métodos de multiplicadores Lagrangianos o modelos adjuntos lineales.

El método 4DVAR es una extensión de 3DVAR que incluye la dimensión temporal en la asimilación de datos, y su objetivo es encontrar la mejor estimación del estado real del sistema (análisis o actualización), considerando tanto las observaciones distribuidas en el tiempo como la dinámica del sistema. De acuerdo con Fischer (2001), las principales diferencias entre los sistemas 3DVAR y 4DVAR son (Beck et al., 2016):

- 1) 4DVAR incluye la integración de un modelo adjunto o de un Modelo lineal tangente durante la minimización.
- 2) En 4DVAR se realizan dos actualizaciones incrementales: una en cada análisis y otra en cada subintervalo. Por ejemplo, un tiempo de análisis de 6 horas puede dividirse en subintervalos de 30 minutos o 1 hora, con asimilación 4DVAR en cada intervalo.
- 3) En el sistema 3DVAR, todas las observaciones para la ventana de asimilación centrada en el tiempo se recogen al mismo tiempo, se comparan con las trayectorias hasta el momento del análisis. En el sistema 4DVAR, las observaciones se dividen en subintervalos y se comparan con las estimaciones, El método 4DVAR es más preciso, teóricamente, que el 3DVAR porque el término se actualiza en cada subintervalo.

La ecuación por minimizar en el caso de 4DVAR, es la siguiente:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b) + \frac{1}{2} \sum_{i=0}^N (H_i(\mathbf{x}_i) - \mathbf{y}_i)^T \mathbf{R}_i^{-1}((\mathbf{x}_i) - \mathbf{y}_i) \dots (2)$$

Con respecto a \mathbf{x} , sujeto a $\mathbf{x}_{i+1} = M_i(\mathbf{x}_i)$. Lo que significa que se tiene que minimizar con respecto a \mathbf{x} sujeto a que tenemos un modelo y el modelo es no lineal.

En el caso de 3DVAR N será igual a cero.

Donde:

- \mathbf{x}^b Estado a priori (background)
- \mathbf{y}_i Observaciones
- H_i Operador de observación
- \mathbf{B} Matriz de covarianza del error en el background
- \mathbf{R}_i Matriz de covarianza del error en las observaciones

Filtro de Kalman

En 1960, Kalman, presenta la solución óptima para la minimización de errores en modelos de dinámica lineal y distribución gaussiana, llamado Filtro de Kalman (KF) (Beck et al., 2016), en este, las observaciones se asimilan secuencialmente, el estado a priori o background viene de un pronóstico previo al análisis y la incertidumbre del background viene de un pronóstico de la incertidumbre en el análisis. El KF es un conjunto de ecuaciones que utilizan el método de mínimos cuadrados para proporcionar una solución óptima de forma recursiva, calculando la solución cada vez que se agrega una observación al sistema (Solera, 2003). En el KF se asume que el modelo y el operador de observación son lineales, además, no se considera el modelo como perfecto, pero se asume que conocemos sus errores y que son insesgados tanto para el estado como para las observaciones (Evensen, 2009).

Existen diferentes versiones del método de Kalman, como lo es su versión extendida, Filtro de Kalman Extendido (EKF), que se aplica a los modelos no lineales, así como el Local Ensemble Transform Kalman Filter (LETKF; Hunt et al., 2007), Además, otra versión del método es el propuesto por Evensen (1994) llamado Filtro de Kalman por Ensamblados (EnKF), este último no necesita la linealización del modelo de pronóstico numérico del tiempo, a

diferencia del método EKF donde si es necesaria la linealización del modelo, además resuelve el problema de los requerimientos computacionales asociados al almacenamiento y posterior integración de la matriz de covarianza del error del estado del modelo. Cada iteración de una aplicación de KF pasa por cuatro pasos:

- 1) Previsión a partir del modelo.
- 2) Cálculo de la matriz de ganancia.
- 3) Cálculo de estimación.
- 4) Análisis.

EKF sigue los mismos pasos que EK sólo se diferencia en que calcula la covarianza de los errores de estimación linealizando el modelo. EnKF consiste en la aplicación simultánea de k ciclos de asimilación de datos. Estos ciclos asimilan las mismas observaciones, pero cada una de ellas recibe una perturbación aleatoria diferente en cada observación (Anderson, 2001). Evensen (1994a, 1994b), menciona que este conjunto de sistemas de asimilación de datos puede utilizarse para estimar la covarianza de los errores, luego de completar k ciclos y obtener las predicciones.

EnKF supone una gran ventaja, a comparación de los anteriores métodos de filtros de Kalman, debido a que, como se mencionó anteriormente, no requiere linealización del modelo (Kalnay, 2003).

A continuación, se muestran las ecuaciones del algoritmo del filtro de Kalman original, conocido como Filtro de Kalman discreto (DKF). Este filtro se utiliza en dos pasos: primero, se realiza un pronóstico del estado futuro tomando en cuenta la información disponible, y luego se genera un pronóstico mejorado del estado, de la tal manera que el error se minimiza estadísticamente (Welch & Bishop, 2001).

Ecuaciones de pronóstico DKF:

$$\mathbf{x}_{t|t-k\Delta t}^- = \mathbf{A}\mathbf{x}_{t-\Delta t|t-(k+1)\Delta t}^+ + \mathbf{B}v_{j,t-\Delta t} \sim N(0, \mathbf{Q}_{t-\Delta t}) \quad \dots (3)$$

$$\mathbf{P}_{x,t|t-k\Delta t}^- = \mathbf{A}\mathbf{P}_{x,t-\Delta t|t-(k+1)\Delta t}^+ + \mathbf{Q}_{t-\Delta t} \dots (4)$$

Ecuaciones de corrección o actualización DKF:

$$\mathbf{K}_{t|t-k\Delta t} = \mathbf{P}_{x,t|t-k\Delta t}^- \mathbf{h}_{t|t-k\Delta t}^T (\mathbf{h}_{t|t-k\Delta t} \mathbf{P}_{x,t|t-k\Delta t}^- \mathbf{h}_{t|t-k\Delta t}^T + \mathbf{R}_t)^{-1} \dots (5)$$

$$\mathbf{x}_{t|t-k\Delta t}^+ = \mathbf{x}_{t|t-k\Delta t}^- + \mathbf{K}_{t|t-k\Delta t} (\mathit{Obs}_t - \mathbf{h}_{t|t-k\Delta t} \mathbf{x}_{t|t-k\Delta t}^-) \dots (6)$$

$$\mathbf{P}_{x,t|t-k\Delta t}^+ = (\mathbf{I} - \mathbf{K}_{t|t-k\Delta t} \mathbf{h}_{t|t-k\Delta t}) \mathbf{P}_{x,t|t-k\Delta t}^- \dots (7)$$

Donde:

$\mathbf{x}_{t|t-k\Delta t}^-$ Estado pronosticado

$\mathbf{P}_{x,t|t-k\Delta t}^-$ Covarianza del error en el background o del pronóstico

$\mathbf{x}_{t|t-k\Delta t}^+$ Estado actualizado

$\mathbf{K}_{t|t-k\Delta t}$ Ganancia de Kalman

$\mathbf{P}_{x,t|t-k\Delta t}^+$ Covarianza del error en la actualización

Obs_t Observaciones

\mathbf{R}_t Covarianza del error en las observaciones

$\mathbf{Q}_{t-\Delta t}$ Covarianza del error en la ecuación de estado

A Matriz que relaciona el estado en el periodo de tiempo previo $t - \Delta t|t - (k + 1)\Delta t$ al estado que ocurre en el momento $t|t - k\Delta t$

B Matriz que relaciona el control opcional de entrada con el estado

Se han realizado comparaciones entre métodos secuenciales y métodos variacionales, ejemplos de ello son:

En 1998, Rabier et al., concluyeron que el método 4DVAR era más eficiente para la ciclogénesis rápida que el método variacional 3DVAR.

En el 2005, Caya et al., compararon el método EnKF y 4DVAR. Utilizando datos sintéticos, asimilaban datos de radar en un modelo de circulación general de la atmósfera, y llegaron a la

conclusión de que el método 4DVAR lograba una mejor previsión de las componentes del viento, mientras que EnKF de la precipitación.

En el 2008, Meng y Zhang, asimilaron datos de presión y perfil vertical del viento, como motivo de comparar EnKF y 3DVAR para el modelo WRF. El ensamble utilizado para los experimentos fue construido para previsiones generadas por diferentes configuraciones de parametrizaciones físicas. Se observó un mejor impacto del método EnKF.

En el 2011, Zhang et al., hicieron una comparación de los métodos EnKF, 3DVAR y 4DVAR en un modelo de área limitada. Informaron un menor error en las predicciones de temperatura y campos de viento para 12, 24 y 36 horas con los métodos EnKF y 4DVAR, así como una mejor precisión con el aumento del horizonte de previsión en EnKF para 48 y 72 horas.

En esta sección se ha presentado el concepto de asimilación de datos, el cual es fundamental para mejorar la precisión de los modelos numéricos del tiempo. Como se mencionó anteriormente, el Filtro de Kalman es un método secuencial de asimilación de datos que combina resultados de modelos numéricos y observaciones para obtener una mejor estimación del estado del sistema. Aunque se utiliza principalmente como método de asimilación de datos, también es comúnmente aplicado en diversos estudios como método de corrección de sesgo de pronósticos meteorológicos numéricos (Homleid 1995; Galanis & Anadranistakis 2002; Galanis et al. 2006; Crochet 2004; Delle Monache et al. 2006; Libionati et al. 2008; Cassola & Burlando 2012; Dorado et al. 2013; Stathopoulos et al. 2013; Pelosi et al. 2017; Righetti et al. 2019; Xu et al. 2021). En esta investigación, se utilizará principalmente para corregir el error sistemático (sesgo) de resultados del sistema SREF. La corrección de sesgo es fundamental para mejorar la precisión de los pronósticos y proporcionar información valiosa para la toma de decisiones, como la gestión de riesgos y la planificación de actividades al aire libre. En la siguiente sección, se hablará sobre la aplicación de Filtro de Kalman como técnica de postprocesamiento estadístico y se describirá su enfoque en la corrección de sesgo de modelos numéricos del tiempo.

1.1.4 Técnicas de postprocesamiento estadístico

Como se mencionó anteriormente, los pronósticos de los modelos de simulación numérica del tiempo presentan errores, producidos por diversos factores. Las técnicas de postprocesamiento estadístico pueden eliminar parte de estos errores. Estas técnicas se pueden clasificar en métodos estadísticos por lotes y métodos adaptativos, estos últimos postprocesan secuencialmente los pronósticos actualizando de manera continua los parámetros de corrección a medida que se dispone de nuevas observaciones (Pelosi et al. 2017). Los métodos adaptativos presentan ciertas ventajas respecto a los métodos estadísticos por lotes, por ejemplo: son computacionalmente menos demandantes, al emplear solo el último conjunto disponible de pronósticos y observaciones en un paso de tiempo dado para actualizar los parámetros de corrección del modelo (Pinson, 2012); son valiosos para sitios donde las estaciones meteorológicas no han operado durante un periodo largo de tiempo y no se dispone de un conjunto de datos de observaciones confiables para utilizar el postprocesamiento por lotes, donde se calibra fuera de línea con un conjunto de entrenamiento de pronósticos y observaciones (Cheng & Steenburgh, 2007; Delle Monache et al., 2006, 2011), además, el procedimiento de actualización de los métodos adaptativos toma en cuenta de manera automática las actualizaciones de los modelos de predicción numérica del tiempo y la no estacionariedad del rendimiento de predicción del modelo respecto a los cambios climáticos estacionales (Persson, 1991).

El Filtro de Kalman es el procedimiento de estimación secuencial estadísticamente óptimo para los sistemas dinámicos. Las observaciones se combinan recursivamente con las previsiones recientes utilizando ponderaciones que minimizan los sesgos correspondientes (Galanis et al., 2006).

Se han implementado procedimientos de Filtro de Kalman a NWP, que en breve serán descritos, para corregir pronósticos de modelos deterministas de variables meteorológicas, como la temperatura. Dicho procedimiento implica la actualización secuencial de los parámetros de las ecuaciones de regresión empleadas para predecir el error de los pronósticos corregidos (predictando) en función de las variables de pronóstico en bruto utilizadas como predictores. Estos procedimientos implementan el mismo esquema de regresión adaptada que

la técnica MOS (Pelosi et al. 2017). El enfoque del Filtro de Kalman soluciona las discrepancias encontradas mediante el uso de la técnica MOS, por ejemplo, en cambios meteorológicos locales de corta duración o las actualizaciones del modelo. Además, la principal ventaja de la aplicación del Filtro de Kalman es la fácil adaptación a cualquier alteración de las observaciones y a su vez puede utilizar series cortas de información (Galanis et al., 2006).

Homleid M. (1995) aplicaron la teoría del Filtro de Kalman para corregir el error sistemático del modelo LAM50. Estudiaron las desviaciones entre los pronósticos de temperatura y las observaciones en diferentes sitios de Noruega. Encontraron que el procedimiento de corrección utilizando Filtro de Kalman redujo el sesgo mensual de los pronósticos de cada estación de observación (240 localizaciones de meteogramas) a valores cercanos a cero.

Galanis G. & Anadranistakis M. (2002) propusieron un filtro de Kalman unidimensional para la corrección de las previsiones de temperatura máxima y mínima a 2 metros al modelo de área limitada Skiron, utilizaron únicamente temperaturas observadas durante un periodo restringido de 7 días. Obtuvieron errores sistemáticos después de aplicar el filtro cercanos a cero, demostrando el mérito del postprocesamiento incluso al aplicar un método sencillo.

Crochet P. (2004) desarrolló un procedimiento de Filtro de Kalman Adaptativo aplicado a las previsiones de temperatura a 2 metros de la superficie, y de la velocidad del viento a 10 metros en Islandia, como objetivo de reducir sesgo sistemático y mejorar la precisión de las previsiones locales derivadas del modelo DMO. El algoritmo de Filtro de Kalman Adaptativo se ejecutó durante el período del 01 de enero de 2001 al 01 de marzo de 2001. Utilizaron observaciones de 11 estaciones sinópticas en Islandia, donde se mostró que para el conjunto de datos probados el método es capaz de eliminar los errores sistemáticos.

Galanis et al., (2006) propusieron una forma de aplicar Filtro de Kalman basada en funciones polinómicas no lineales. Utilizaron salidas de modelos NWP y observaciones de temperatura y velocidad del viento de dos lugares al sur de Europa para diferentes periodos de tiempo. Para ello, aplicaron un Filtro de Kalman a un conjunto de datos de un año donde obtuvieron

una mejora directa del modelo. En general, la metodología propuesta mostró un buen rendimiento.

Delle Monache et al., (2006) utilizaron el método de corrección de sesgo de postprocesamiento del predictor del Filtro de Kalman a 12 miembros de un modelo de calidad del aire y a su media corregida, utilizando datos de cinco sitios donde se realizan observaciones de Ozono en Canadá en un periodo que comprende del 11 al 15 de agosto del 2004. Realizaron la corrección de sesgo con Filtro de Kalman a cada uno de los miembros y calcularon la media del ensamble corregida, y posteriormente a esa media también le implementaron el Filtro para corregirla. En dicho estudio se identificó una reducción tanto en los errores sistemáticos (sesgo) como en los no sistemáticos; siendo la corrección aplicada a la media del ensamble corregida la que muestra resultados más satisfactorios.

Libionati et al., (2008) emplearon un Filtro de Kalman a las previsiones de temperatura a 2 metros emitidas en el año 2003 por el European Centre for Medium-Range Weather Forecast (ECMWF) en 12 estaciones sinópticas situadas en Portugal. Encontraron que la metodología propuesta permite una corrección objetiva de las previsiones de temperatura a 2 metros, reduciendo el sesgo a valores cercanos a cero, y mejorando el error cuadrático medio desde el 10% hasta más del 70%, con respecto a las previsiones brutas del ECMWF.

Cassola F. & Burlando M. (2012) implementaron el Filtro de Kalman a salidas proporcionadas por el modelo BOLAM para mejorar localmente las previsiones de velocidad del viento. Utilizaron un conjunto de datos observados de velocidad del viento de un periodo de 2 años para dos estaciones anemométricas ubicadas en el este de Liguria (Italia). Encontraron que, ajustando el paso temporal y el horizonte de previsión del filtro, la metodología es capaz de proporcionar una mejora significativa de la previsión con respecto a la salida del modelo en bruto. Después de dos años de prueba, la metodología se mantuvo estable y el porcentaje de error continuaba siendo muy bajo.

Dorado et al., (2013) implementaron la metodología de Filtro de Kalman como método de ajuste a las salidas del modelo GFS para el pronóstico de temperatura máximas y mínimas en

27 estaciones ubicadas en Colombia. La implementación del Filtro la realizaron en dos partes, una como método de ajuste y otra como modelo de pronóstico. Encontraron resultados satisfactorios, mejorando significativamente las series del orden de un 90%.

Stathopoulos et al., (2013) proponen una combinación de modelos físicos y estadísticos para alcanzar una estimación local precisa de las condiciones de viento sobre las islas de Crete y Kefalonia en Grecia. Ellos emplearon los modelos SKIRON (modelo no hidrostático) y el sistema RAMS (fusión entre un modelo de nubes no hidrostático y un modelo de mesoescala hidrostático) para proporcionar pronósticos de viento local en los sitios de interés y se utilizaron postprocesos estadísticos basados en Filtro de Kalman y Kolmogorov-Zurbenko, para reducir los sesgos sistemáticos. Encontraron que el filtro de Kalman mejora significativamente el desempeño de los pronósticos al eliminar casi por completo el sesgo.

Pelosi et al., (2017) propusieron e implementaron un nuevo enfoque de Filtrado de Kalman para la corrección adaptativa del error de pronóstico en salidas de pronósticos por ensamble para Campania, en el sur de Italia. Aplicaron los métodos AMOS (utilizando la media del ensamble) y AEMOS (explotando completamente el contenido de los miembros del ensamble para actualizar los parámetros de la ecuación) para el postprocesamiento de los pronósticos por ensamble COSMO-LEPS. Utilizaron dos años de observaciones meteorológicas (2014 y 2015) de las variables de temperatura a 2 m y velocidad del viento de 18 estaciones meteorológicas automáticas para la verificación de los pronósticos. Mostraron que el método AEMOS supera al método AMOS.

Righetti et al., (2019) realizaron la corrección del error sistemático en la predicción numérica de la temperatura del aire mediante la técnica de Regresión Adaptada basada en el Filtro de Kalman. Corrigieron los pronósticos de temperatura mínima, máxima y de las 21 horas en Argentina, del modelo global GFS con datos desde mayo del 2016, comparándolas con estaciones meteorológicas para Argentina. La implementación de la metodología resultó en resultados favorables en la corrección de las 3 variables analizadas, la metodología RAFK logra valores del BIAS en torno a cero, y RMSE entre 1.5 y 2 °C.

Xu et al., (2021) aplicaron un enfoque de Filtro de Kalman de corrección de sesgo de postprocesamiento a pronósticos de velocidad del viento de 72 horas a 70 metros de altura del modelo WRF en el parque eólico de Zhangbei durante un periodo de más de dos años. Encontraron que el Filtro de Kalman no solo reduce los errores sistemáticos de modelo, sino que también reduce significativamente los errores aleatorios. Los resultados muestran una disminución de RMSE en un 16 % (de 3.58 a 3.01 ms^{-1}), MAE en un 14% (de 2.71 a 2.34 ms^{-1}), sesgo de 0.22 a -0.19 ms^{-1} y mejora de la correlación de 0.58 a 0.66. Mencionan que los errores están dominados por el componente aleatorio.

Como se ha demostrado en los estudios previamente mencionados, el Filtro de Kalman Adaptativo ha sido utilizado para corregir el error sistemático en los modelos de simulación numérica del tiempo, y los resultados obtenidos han sido favorables. Por consiguiente, la importancia de la metodología radica en la mejora de precisión de las predicciones, lo que contribuye significativamente en el desarrollo de pronósticos meteorológicos más precisos, en beneficio de los usuarios finales de la información.

Planteamiento del problema

Uno de los principales desafíos a los que se enfrentan los meteorólogos y/o los consultores del pronóstico del tiempo, es la incertidumbre que existe entre los datos pronosticados y los observados. Según Morss et al. (2008), una de las causas de la incertidumbre presente en los pronósticos corresponde a que la atmósfera es un sistema altamente no-lineal, por lo que las predicciones son aproximadas y no exactas.

Los modelos numéricos presentan errores al pronosticar variables meteorológicas, principalmente las cercanas a superficie, debido a las limitaciones en cuanto a representación de procesos físicos por parte de las parametrizaciones, además a la inhabilidad para resolver fenómenos de pequeña escala (Righetti et al., 2019). Estos errores poseen tanto una parte sistemática (sesgo) como una parte aleatoria (random). La parte sistemática es comúnmente removida mediante métodos estadísticos (Kalnay, 2003); mientras que la parte aleatoria suele ser difícil de estimar, debido a la complejidad en separar la inexactitud del modelo y el estado inicial del error (Jung et al., 2005).

Estos errores de predicciones para el estado del tiempo atmosférico pueden tener consecuencias negativas para la sociedad, debido a que malos pronósticos pueden traer consigo pérdidas económicas, en actividades como la agricultura, pesca, así como afectaciones en los ámbitos de salud, recreación, seguridad, entre otros.

Objetivos

1.2.1 Objetivo general

- Implementar una metodología de corrección de sesgo, basada en filtro de Kalman Adaptativo, para disminuir las diferencias entre resultados de simulaciones numéricas del tiempo y observaciones meteorológicas en la variable de temperatura (máxima y mínima) y precipitación en el territorio mexicano.

1.2.2 Objetivos específicos

- Recopilar y organizar datos provenientes de salidas del sistema Short-Range Ensemble Forecast (SREF) para conformar un arreglo de datos.
- Obtención de observaciones meteorológicas de estaciones climáticas convencionales.
- Identificación de las diferencias entre los pronósticos de SREF y observaciones meteorológicas para evaluar la precisión y sesgo del modelo.
- Aplicación e implementación de Filtro de Kalman Adaptativo para corregir el sesgo del sistema SREF.
- Realizar un análisis estadístico de los resultados obtenidos al aplicar Filtro de Kalman Adaptativo para identificar la reducción del sesgo.
- Documentar los resultados obtenidos, así como los aspectos científicos y tecnológicos relacionados con el uso de modelos numéricos, pronóstico meteorológico y métodos estadísticos para corrección de sesgo.

Descripción de la región de estudio

En esta sección, se brinda una descripción de la región de estudio, con el propósito de comprender los factores que influyen el comportamiento de las variables meteorológicas analizadas en este estudio.

Ubicación de Estaciones Climáticas Convencionales



Figura 2: Distribución de estaciones climáticas convencionales, a lo largo del territorio mexicano. Elaboración propia mediante QGIS.

La República Mexicana se encuentra situada entre los paralelos $14^{\circ} 28'$ y $32^{\circ} 43' 35''$ de latitud norte. De acuerdo con López de Llergo (2003), la República Mexicana se encuentra en la zona subecuatorial, tropical y subtropical del Hemisferio Norte, con latitudes extremas de 14° y 32° N. La altitud tiene mayor influencia en la distribución de las temperaturas que la latitud.

Se identifica una línea llamada Trópico Climatológico que delimita climas cálidos y

semicálidos al sur, y una estación fría bien definida al norte. Las curvas de altitud de 1,000 m, 1,500 m y 2,500 m separan distintos tipos de climas. Los vientos predominantes son del noreste debido a los alisios del Hemisferio Norte. En invierno, la región noroeste de la Altiplanicie y el norte de Baja California reciben brisas del oeste. Los alisios cargados de humedad chocan contra las vertientes este de la Sierra Madre Oriental y norte de las Sierras Septentrionales de Chiapas, generando precipitaciones intensas. La pluviosidad máxima se encuentra entre los 100 y 1,500 m de altitud. Al pasar las sierras mencionadas, los vientos generan una transición abrupta de climas húmedos a esteparios y desérticos (López de Llergo, 2003).

En las regiones al oeste de la Sierra Madre Oriental y al sur de las Sierras Septentrionales de Chiapas, las lluvias son proporcionadas por los ciclones antillanos. Estos ciclones se forman debido a las diferencias de presión entre el norte del Golfo de México y los mares de las Antillas y el Caribe. Provocan precipitaciones en las vertientes meridionales de la Sierra Madre del Sur, el Eje Volcánico y las Sierras de Zacatecas, generando climas húmedos en las vertientes exteriores del sur de la Sierra Madre Occidental y la región este de la Sierra Madre de Chiapas (López de Llergo, 2003).

En invierno, soplan ondas frías conocidas como "nortes" desde el Golfo de Tehuantepec y los mares de las Antillas y Caribe. Estos vientos son fríos y secos en la mayoría del país, excepto en ciertas áreas donde llegan después de atravesar el Golfo de México (López de Llergo, 2003).

La parte norte de la Altiplanicie y Baja California tienen climas esteparios y desérticos, aunque algunos ciclones atraviesan estas zonas después de dejar su humedad en las vertientes meridionales de las sierras mencionadas (López de Llergo, 2003).

De acuerdo con el Nuevo atlas nacional de México, se identifican siete zonas térmicas en el país: cálida, semicálida, templada, semifría, fría y muy fría. La región más cálida se encuentra en la llanura costera del Golfo de México, Tabasco, Campeche, la Península de Yucatán y una

franja en la vertiente del Océano Pacífico. La región semicálida está en las laderas bajas de las principales cadenas montañosas, entre 500 y 1,000 m de altitud. La región templada abarca las laderas altas y mesetas con temperaturas por debajo de 18°C, generalmente sobre los 1,000 m de altitud (López de Llergo, 2003).



Figura 3: Temperatura media anual en el territorio mexicano. Obtenido de Coll-Hurtado, 1990.

El clima en México muestra variaciones significativas a lo largo del año y en diferentes regiones. Durante el ciclo anual, se presentan invasiones de aire tropical que generan ondas de calor en distintos momentos y lugares del país. Estas ondas de calor se observan en abril y mayo en el centro y sur de México, en junio y julio al norte de la Altiplanicie mexicana, y en agosto y septiembre en el noreste y la Península de Baja California (Coll-Hurtado, 1990).

Además del ciclo anual, el clima mexicano también sigue un ciclo diario debido a la sucesión del día y la noche. Las temperaturas máximas diarias suelen ocurrir después del mediodía, aproximadamente tres horas más tarde, mientras que las temperaturas mínimas se registran poco después de la salida del sol. La hora de la temperatura mínima puede variar según la hora de la salida del sol (Coll-Hurtado, 1990).

En cuanto a la precipitación, se observan notables diferencias en su distribución a lo largo del territorio mexicano. En la vertiente del Golfo de México, se destaca una zona especialmente húmeda al sur del paralelo 22°N, donde se encuentran las pendientes montañosas del centro y sur del país. Estas áreas están directamente expuestas a los vientos húmedos provenientes del Golfo de México, lo cual genera precipitaciones abundantes al encontrarse con las barreras montañosas (Coll-Hurtado, 1990).

Dentro de la vertiente del Golfo de México, se identifican cuatro áreas con una precipitación anual superior a 2,500 mm. Estas áreas incluyen las Sierras de Teziutlán y Zacapoaxtla en Puebla, las Sierras de Ixtlán y Mixes en Oaxaca, la Sierra de los Tuxtlas en Veracruz y las Montañas del Norte de Chiapas. Por otro lado, en la vertiente del Océano Pacífico, las cantidades de lluvia son más moderadas, registrando valores superiores a 2,000 mm únicamente en la porción sur de la Sierra Madre de Chiapas (Coll-Hurtado, 1990).

En términos generales, las regiones montañosas de México reciben un promedio de precipitaciones que oscila entre los 800 y 1,500 mm al año, mientras que en el norte y noroeste del país las lluvias son escasas, con cifras inferiores a los 600 mm anuales. En las zonas más áridas, como la Península de Baja California y la desembocadura del río Colorado, las precipitaciones son aún más limitadas, llegando a ser inferiores a los 125 mm al año (Coll-Hurtado, 1990).

En cuanto a los ciclones tropicales, son fenómenos hidrometeorológicos que ocasionan precipitaciones en México. Entre 1970 y 2017, se han contabilizado un total de 236 ciclones tropicales, de los cuales 22 alcanzaron la categoría de huracanes intensos con velocidades máximas sostenidas superiores a 178 km/h. Estos huracanes no solo representan un desafío

para las costas mexicanas, sino que también transportan humedad hacia el interior del país (CONAGUA, 2019).

En cuanto a las aguas superficiales, México cuenta con una extensa y diversa red hidrográfica que abarca aproximadamente 633 mil kilómetros de longitud. Destacan 51 ríos principales que desempeñan un papel fundamental en el sistema hídrico del país, siendo responsables del 87% del escurrimiento superficial y cubriendo alrededor del 65% de la superficie nacional. Entre las cuencas más destacadas se encuentran los ríos Bravo y Balsas, conocidos por su superficie, así como los ríos Bravo y Grijalva-Usumacinta que ocupan un lugar destacado en términos de longitud. Otros ríos importantes son el Lerma y el Nazas-Aguanaval, que forman parte de la vertiente interior y desempeñan un papel vital en el sistema hídrico de sus respectivas regiones (CONAGUA, 2019).

Es importante destacar que este estudio se basa en el análisis de 1570 estaciones climáticas convencionales. La Figura 2 muestra la ubicación de las estaciones climáticas convencionales ubicadas dentro de la zona de estudio, al momento se tiene una serie de datos de temperatura máxima, mínima y precipitación, a partir de marzo del 2020 a abril del 2021. Además, se cuenta con un almacenamiento de resultados del sistema SREF, a partir del 01 de abril del 2020 al 21 de mayo del 2022. Se realizó la extracción de los datos, utilizando la media del ensamble, correspondientes de SREF utilizando las coordenadas de 1570 estaciones climáticas convencionales localizadas en nuestra zona de interés.

Justificación

Actualmente existen diversos modelos de simulación numérica del tiempo que nos ayudan a estimar la evolución de los flujos atmosféricos, sin embargo, estos presentan errores en los pronósticos de las variables meteorológicas, principalmente en las cercanas a la superficie. De acuerdo con Bjerknes (1911) la predicción meteorológica es un problema de condiciones iniciales. Teniendo las condiciones iniciales y de frontera apropiadas se puede representar con cierta precisión la evolución del estado atmosférico. A su vez, las técnicas de postprocesamiento estadístico pueden ser aplicadas para corregir parcialmente el impacto de los errores de modelos NWP, y, por lo tanto, mejorar el rendimiento de las salidas (Vannitsem, 2008). Dentro del postprocesamiento estadístico se encuentra el método adaptativo, que postprocesa secuencialmente los pronósticos mediante la actualización continua de los parámetros de corrección a medida que se dispone de observaciones (Pelosi et al., 2017).

Este tema de investigación surge de la necesidad de mejorar la calidad de las salidas de los modelos numéricos mediante la implementación de una metodología de corrección de sesgo que contribuya a elevar la efectividad de los pronósticos meteorológicos.

Se espera que el proceso de identificación y corrección de sesgo mediante la implementación de un Filtro de Kalman Adaptativo en las salidas del sistema Short-Range Ensemble Forecast (SREF) resulte en la reducción del error entre los valores simulados por dicho ensamble y los valores observados en estaciones meteorológicas, con el fin de obtener pronósticos numéricos más acertados para la república mexicana, para beneficio de la sociedad.

Alcances y limitaciones

El alcance principal de este trabajo fue implementar un método para corregir el sesgo en los resultados de simulaciones numéricas del tiempo del modelo SREF, específicamente la aplicación de Filtro de Kalman Adaptativo. De esta manera, se busca disminuir las diferencias que existen entre los pronósticos del modelo SREF y observaciones de las variables meteorológicas de temperatura máxima, mínima y precipitación.

Sin embargo, durante la realización de esta investigación se presentaron ciertas limitaciones, una de ellas es contar una óptima densidad de estaciones que proporcionen información de calidad. Además, se requirió de un equipo de cómputo con suficiente capacidad de procesamiento y almacenamiento para respaldar las simulaciones del sistema SREF. Por lo cual, se dependió de una estación de trabajo para almacenar los pronósticos numéricos.

En cuanto a la aplicación del Filtro de Kalman Adaptativo en variables cuyo comportamiento varía significativamente tanto espacial como temporalmente, como ocurre con la precipitación, se ha observado que las estimaciones no son precisas. Por lo tanto, se recomienda que en investigaciones futuras se identifique el método óptimo para corregir la estimación de esta variable en México.

Además de las limitaciones mencionadas anteriormente, también se consideró la posibilidad de realizar la corrección de sesgo en tiempo real mediante la utilización de Estaciones Meteorológicas Automáticas (EMAs), aunque no es el objetivo del trabajo. Estas estaciones, equipadas con sensores y dispositivos electrónicos, proporcionan mediciones meteorológicas continuas y en tiempo real, lo que las convierte en una opción atractiva para corregir el sesgo de los pronósticos meteorológicos numéricos en tiempo real.

Sin embargo, a pesar de considerar las EMAs como una opción potencial, se encontraron desafíos significativos para acceder a la información en tiempo real de estas estaciones. Estos desafíos se resumen en dificultades para el acceso de los datos de manera oportuna.

Aunque no se pudo acceder a los datos de las EMAs en tiempo real para realizar la corrección de sesgo, es importante destacar la relevancia de este enfoque. La utilización de EMAs en futuras investigaciones podría permitir una corrección actualizada de los pronósticos meteorológicos, lo que mejoraría la calidad de los resultados y la capacidad de anticiparse a cambios repentinos en las condiciones atmosféricas.

Hipótesis

La implementación de una metodología de corrección de sesgo, basada en la aplicación de Filtro de Kalman Adaptativo, reduce las diferencias entre resultados de simulaciones numéricas del tiempo del modelo SREF y observaciones puntuales de las variables meteorológicas de temperatura máxima, temperatura mínima y precipitación en la República Mexicana.

Equipo de cómputo y lenguaje de programación

Durante el desarrollo de esta investigación, se utilizaron dos sistemas operativos, uno para la recolección de la información del sistema SREF y otro para el postprocesamiento de los datos. Los datos provenientes del sistema SREF se descargaron y almacenaron en el sistema operativo LINUX, mientras que el sistema operativo Windows 10 se utilizó para el postprocesamiento de los mismos.

Para desarrollar el algoritmo de corrección de sesgo basado en Filtro de Kalman Adaptativo, se empleó el lenguaje de programación Python, utilizando Spyder como su IDE (Integrated Development Environment). Se programaron las ecuaciones del filtro y se calcularon diversas métricas estadísticas en Python para identificar el sesgo del modelo y para evaluar el rendimiento de la corrección mediante la aplicación del Filtro de Kalman Adaptativo.

Python es un lenguaje de programación de alto nivel, interpretado y multipropósito, utilizado en una amplia variedad de campos, desde la ciencia de datos hasta el desarrollo web. Es conocido por su facilidad de lectura y escritura, sintaxis clara y expresiva, y una amplia variedad de librerías especializadas en diversas áreas (Python Software Foundation, 2023). Spyder, por otro lado, es un entorno de desarrollo integrado (IDE, por sus siglas en inglés) diseñado específicamente para programar en Python. Permite la edición de código, la depuración, la ejecución de scripts, la creación de gráficos y mucho más. Además, Spyder viene preinstalado con muchas de las bibliotecas científicas más comunes utilizadas en

Python, lo que lo convierte en una opción popular para la ciencia de datos y la investigación (Spyder IDE, 2023).

En el próximo capítulo, se describirá de manera general las métricas estadísticas utilizadas y se dará una descripción detallada del Filtro de Kalman Adaptativo utilizado para corregir el sesgo del modelo SREF.

Capítulo 2. Metodología

Métricas de evaluación

En este trabajo, se utilizaron un conjunto de métricas estadísticas para medir la precisión del pronóstico y evaluar el rendimiento de los pronósticos del sistema SREF y de la corrección a través de Filtro de Kalman Adaptativo. Algunas de las métricas estadísticas que se utilizarán son: la raíz del error cuadrático medio (RMSE), el error absoluto medio (MAE), el coeficiente de correlación de person (COE) y el sesgo (BIAS) definidos por Pielke (1984) y Stauffer y Seaman (1990).

El cálculo de la precisión se basa en el RMSE, también conocido como raíz del error cuadrático medio, que se define como (Pielke, 1984):

$$RMSE = \sqrt{\sum_{i=1}^N \frac{(\phi_i - \phi_{iobs})^2}{N}} \dots (8)$$

El RMSE proporciona una medida de las diferencias entre los valores pronosticados y los observados, teniendo en cuenta la magnitud del error cuadrático.

Además, se empleó el MAE, que se define como (Stauffer y Seaman, 1990):

$$MAE = \sum_{i=1}^N \frac{|\phi_i - \phi_{iobs}|}{N} \dots (9)$$

El MAE proporciona información sobre el error absoluto promedio entre los valores pronosticados y observados.

Por otra parte, el sesgo (BIAS) nos proporciona información sobre la tendencia del modelo a sobreestimar o subestimar una variable, nos cuantifica el error sistemático del modelo, Pielke (1984) define el BIAS según:

$$BIAS = \sum_{i=1}^N \frac{(\phi_i - \phi_{iobs})}{N} \dots (11)$$

Para evaluar la relación lineal entre los valores pronosticados y observados, se utilizó el coeficiente de correlación (Dorado, 2013). El coeficiente de correlación cuantifica la fuerza y dirección de esta relación. Se utilizó el coeficiente de correlación de Pearson para este análisis.

$$COE = \frac{\sum_{i=1}^N [(\phi_{iobs} - \overline{\phi_{iobs}}) * (\phi_i - \overline{\phi_i})]}{\sqrt{\sum_{i=1}^N (\phi_{iobs} - \overline{\phi_{iobs}})^2 * \sum_{i=1}^N (\phi_i - \overline{\phi_i})^2}} \dots (12)$$

Donde:

ϕ_i es el valor pronosticado para la celda i

$\overline{\phi_i}$ es el promedio del valor pronosticado por el modelo

ϕ_{iobs} es el valor observado para la celda i

$\overline{\phi_{iobs}}$ es el promedio del valor observado por la estación meteorológica

N es el número de valores analizados

Con esta base, en secciones posteriores se emplearán los métodos estadísticos mencionados, para cuantificar tanto el sesgo como las diferencias presentes en los resultados del sistema SREF en relación a las observaciones de las variables de temperatura máxima, temperatura mínima y precipitación.

Método de corrección basado en filtros de Kalman

A continuación, se describe la estructura general del procedimiento de Filtro de Kalman Adaptado que se aplicó para realizar la corrección de sesgo en las salidas del sistema SREF. Para esto, utilizamos la variable de pronóstico en bruto como predictor, y el error de pronóstico como predictando, esto de manera análoga al esquema de corrección MOS. Se examinan los pronósticos brutos agregados en el paso de tiempo Δt , con tiempos de entrega $k\Delta t$ con $k = 1, \dots, r$, donde $r\Delta t$ es el rango de pronóstico (Pelosi et al, 2017).

El error de pronóstico con respecto a la observación está definido como:

$$y_{t|t-k\Delta t} = z_{t|t-k\Delta t} - Obs_t \dots (13)$$

Donde $z_{t|t-k\Delta t}$ es el pronóstico en bruto de una variable meteorológica y Obs_t es la observación registrada.

La manera de explicar el comportamiento del error de pronóstico, se establece mediante una función polinómica de la variable de pronóstico $z_{t|t-k\Delta t}$:

$$y_{t|t-k\Delta t} = x_{0,t} + x_{1,t}z_{t|t-k\Delta t} + x_{2,t}z_{t|t-k\Delta t}^2 + \dots + x_{p,t}z_{t|t-k\Delta t}^p + \eta_t \dots (14)$$

Donde η_t es el ruido (error) y $x_{j,t}$ denota el Jth-esimo coeficiente de la función polinómica de grado P .

Ahora, se supone que los coeficientes de la función polinómica dependen del tiempo, y evolucionan en el tiempo de acuerdo con un paso aleatorio, y está representado por:

$$x_{j,t|t-k\Delta t} = x_{j,t-\Delta t|t-(k+1)\Delta t} + v_{j,t-\Delta t} \dots (15)$$

Donde $v_{j,t-\Delta t}$ representa el ruido de la estimación de los coeficientes de regresión.

La ecuación del sistema o estado (14) y la de observación (15) pueden escribirse de forma matricial de la siguiente manera:

Ecuación de estado en forma matricial:

$$\mathbf{x}_{t|t-k\Delta t} = \mathbf{x}_{t-\Delta t|t-(k+1)\Delta t} + \mathbf{v}_{j,t-\Delta t} \sim N(0, \mathbf{Q}_{t-\Delta t}) \dots (16)$$

Ecuación de observación en forma matricial:

$$y_{t|t-k\Delta t} = \mathbf{h}_{t|t-k\Delta t} \mathbf{x}_{t|t-k\Delta t} + \eta_t \sim N(0, R_t) \dots (17)$$

Donde:

$\mathbf{v}_{j,t-\Delta t} \sim N(0, \mathbf{Q}_{t-\Delta t})$ es el ruido en la estimación de los coeficientes de regresión y se asume una distribución normal con media 0 y covarianza $\mathbf{Q}_{t-\Delta t}$.

$\eta_t \sim N(0, R_t)$ es el ruido en las variables pronosticadas con distribución normal y covarianza R_t .

El operador de observación \mathbf{h} se define como: $h_{t|t-k\Delta t} = [1, z_{t|t-k\Delta t}, z_{t|t-k\Delta t}^2, \dots, z_{t|t-k\Delta t}^p]$

La covarianza del error de la ecuación de estado $\mathbf{Q}_{t-\Delta t}$ y la covarianza del error de la ecuación de observación R_t , fueron calculadas con las ecuaciones (18) y (19) respectivamente, de acuerdo con (Galanis & Anadranistokis , 2002), donde proponen estimar adaptativamente las varianzas de ruidos a partir de los datos de los últimos 7 días de los coeficientes de regresión predichos y los errores de pronóstico:

$$\mathbf{Q}(t) = \frac{1}{6} * \sum_{i=0}^6 ((x_{t-i|t-k\Delta t-i} - x_{t-\Delta t-i|t-(k+1)\Delta t-i}) - ((\frac{\sum_{i=0}^6 (x_{t-i|t-k\Delta t-i} - x_{t-\Delta t-i|t-(k+1)\Delta t-i})}{7}))^2) \dots (18)$$

$$R(t) = \frac{1}{6} * \sum_{i=0}^6 ((y_{t-i|t-k\Delta t-i} - h_{t-i|t-k\Delta t-i} * x_{t-\Delta t-i|t-(k+1)\Delta t-i}) - ((\frac{\sum_{i=0}^6 (y_{t-i|t-k\Delta t-i} - h_{t-i|t-k\Delta t-i} * x_{t-\Delta t-i|t-(k+1)\Delta t-i})}{7}))^2) \dots (19)$$

La estimación previa de los coeficientes de regresión $\mathbf{x}_{t|t-k\Delta t}^-$ y su matriz de covarianza del error $\mathbf{P}_{x,t|t-k\Delta t}^-$ se definen de la siguiente manera:

$$\mathbf{x}_{t|t-k\Delta t}^- = \mathbf{x}_{t-\Delta t|t-(k+1)\Delta t}^+ \dots (20)$$

$$\mathbf{P}_{x,t|t-k\Delta t}^- = \mathbf{P}_{x,t-\Delta t|t-(k+1)\Delta t}^+ + \mathbf{Q}_{t-\Delta t} \dots (21)$$

Dados los pronósticos brutos disponibles en el tiempo t, el error de pronóstico se estima de acuerdo con:

$$\hat{y}_{t|t-k\Delta t} = \mathbf{h}_{t|t-k\Delta t} \mathbf{x}_{t|t-k\Delta t}^- \dots (22)$$

A medida que tenemos una observación disponible para el tiempo t, la innovación del error con respecto al error predicho se calcula mediante:

$$\tilde{y}_{t|t-k\Delta t} = y_{t|t-k\Delta t} - \hat{y}_{t|t-k\Delta t} \dots (23)$$

La predicción previa actual se combina con la observación para proporcionar una estimación actualizada de los coeficientes de regresión $\mathbf{x}_{t|t-k\Delta t}^+$ (24) y de la covarianza del error $\mathbf{P}_{x,t|t-k\Delta t}^+$ (25) de la siguiente forma:

$$\mathbf{x}_{t|t-k\Delta t}^+ = \mathbf{x}_{t|t-k\Delta t}^- + \mathbf{K}_{t|t-k\Delta t} \tilde{y}_{t|t-k\Delta t} \dots (24)$$

$$\mathbf{P}_{x,t|t-k\Delta t}^+ = (\mathbf{I} - \mathbf{K}_{t|t-k\Delta t} \mathbf{h}_{t|t-k\Delta t}) \mathbf{P}_{x,t|t-k\Delta t}^- \dots (25)$$

$$\mathbf{S}_{t|t-k\Delta t} = \mathbf{h}_{t|t-k\Delta t} \mathbf{P}_{x,t|t-k\Delta t}^- \mathbf{h}_{t|t-k\Delta t}^T + R_t \dots (26)$$

$$\mathbf{K}_{t|t-k\Delta t} = \mathbf{P}_{x,t|t-k\Delta t}^- \mathbf{h}_{t|t-k\Delta t}^T \mathbf{S}_{t|t-k\Delta t}^{-1} \dots (27)$$

Donde $\mathbf{K}_{t|t-k\Delta t}$ es la ganancia de Kalman y $S_{t|t-k\Delta t}$ es la varianza de las innovaciones.

Una vez se estiman los nuevos coeficientes $\mathbf{x}_{t|t-k\Delta t}^+$, después de que se ha completado la actualización que se hizo con base al análisis de las observaciones en el tiempo t y los pronósticos emitidos en el momento $t - k\Delta t$ válido para el tiempo t . El pronóstico corregido para el siguiente paso de tiempo en anticipación $z_{t+k\Delta t|t}$ (*corregida*), se calcula haciendo uso de los coeficientes de correlación actualizados $\mathbf{x}_{t|t-k\Delta t}^+$ de la siguiente manera (Pelosi, 2017):

$$z_{t+k\Delta t|t}(\text{corregida}) = z_{t+k\Delta t|t} - y_{t+k\Delta t|t}^+ \dots (28)$$

Donde: $y_{t+k\Delta t|t}^+ = \mathbf{h}_{t+k\Delta t|t} \mathbf{x}_{t|t-k\Delta t}^+$

Por lo que, para corregir el pronóstico meteorológico numérico para el día de mañana $z_{t+k\Delta t|t}$, se utilizan la ecuación de predicción (22) en el tiempo $t + k\Delta t|t$, que se basa en los coeficientes actualizados calculados con las ecuaciones de corrección (24) en el tiempo anterior $t - k\Delta t$. El pronóstico corregido se emite para el tiempo posterior $t + k\Delta t$, pero es válido para el tiempo t . En la [Figura 10](#) se muestra un diagrama que resume la ecuaciones descritas en esta sección, en dicha figura se observa la estimaciones iniciales necesarias para inicializar el algoritmo, además del tiempo de aplicación de cada conjunto de ecuaciones.

Para la primera estimación se necesitan valores iniciales de coeficientes de regresión $\mathbf{x}_{t|t-k\Delta t}^-$ y de la covarianza de su error $\mathbf{P}_{x,t|t-k\Delta t}^-$, sin embargo, se menciona que su impacto tiende a desaparecer rápidamente después de las primeras actualizaciones y convergen rápidamente a sus valores verdaderos (Pelosi, 2017). Por otro lado, las covarianzas del ruido $\mathbf{Q}_{t-\Delta t}$ y R_t son más importantes ya que impactan el desempeño del filtro (Galanis & Anadranistokis, 2002).

Las ecuaciones mencionadas en esta sección fueron discretizadas en el lenguaje de programación Python y posteriormente se aplicaron en casos de estudio. Se proporcionarán más detalles sobre esta implementación en secciones posteriores.

Capítulo 3. Desarrollo

Datos meteorológicos utilizados: observaciones y pronósticos

En esta sección, se presentará una descripción detallada de los datos utilizados en este trabajo, destacando su relevancia en el proceso de corrección de sesgo y en la aplicación del Filtro de Kalman Adaptativo. Se utilizaron como datos de pronósticos los resultados de simulaciones numéricas del tiempo del sistema SREF, mientras que como observaciones se tomaron datos de estaciones climáticas convencionales distribuidas en todo el país. Estas observaciones fueron fundamentales en la metodología del filtro, ya que se utilizaron para corregir el sesgo de las variables de temperatura máxima, temperatura mínima y precipitación. En las siguientes secciones, se explicará en detalle la metodología utilizada y cómo se emplearon estos datos en la aplicación del filtro.

A continuación, se procede con la descripción de los datos:

2.3.1 Datos observados de estaciones climáticas convencionales

Se disponen de observaciones de temperatura máxima (°C), temperatura mínima (°C) y precipitación acumulada (mm) diarias, de una red de 1570 estaciones climáticas convencionales, de marzo del 2020 a abril del 2021, procedente de la base de datos del CLImate COMputing project (CLICOM), administrada por el Servicio Meteorológico Nacional (SMN).

De acuerdo con Quintas (2000), las estaciones climatológicas son asistidas por un observador que las visita alrededor de las 8:00 am para realizar las mediciones y registrar las variables según un formato determinado.

En estas estaciones se registran 5 variables cuantitativas y 5 cualitativas. Para nuestro interés, las variables de temperatura máxima, temperatura mínima y precipitación se registran de la siguiente manera:

La temperatura máxima se mide utilizando un termómetro de máxima y se registra en grados Celsius con un dígito decimal. La lectura se realiza alrededor de las 8:00 am y representa la temperatura más alta ocurrida posiblemente entre las 11:00 am y las 5:00 pm del día anterior.

La temperatura mínima se mide utilizando un termómetro de mínima y se registra en grados Celsius con un dígito decimal. La lectura también se realiza alrededor de las 8:00 am y corresponde a la temperatura más baja ocurrida durante la madrugada del día de la lectura.

En cuanto a la precipitación, se registra la lámina de lluvia acumulada en las 24 horas anteriores a las 8:00 am del día de la medición. Se mide en milímetros con un dígito decimal. La cantidad de lluvia registrada puede corresponder a la lluvia caída la tarde anterior o, en el caso de lluvias durante la madrugada, podría representar la lámina de lluvia acumulada durante esa madrugada.

Es importante destacar que la calidad de las mediciones de estas variables depende del estado de calibración de los instrumentos utilizados y de la habilidad del observador (Quintas, 2000). Además, el SMN realiza un control de calidad inicial de la base de datos del CLICOM (CICESE, 2013).

2.3.2 Datos pronosticados por el sistema SREF

Se utilizaron datos del sistema SREF, los cuales se han descargado y almacenado, se cuenta con periodo desde abril del 2020 a junio del 2020. Aunque se dispone de datos de diversas variables meteorológicas, para los fines de este trabajo se seleccionaron únicamente las variables de temperatura y precipitación.

Los datos del sistema SREF se encuentran registrados en hora Zulu y tienen un espaciamiento temporal de 3 horas. Para calcular la temperatura máxima y mínima, se consideró el intervalo de tiempo comprendido entre las 15Z (09:00 am, hora local) y las 12Z del día siguiente (06:00 am, hora local) (consultar tabla 1 para una mejor comprensión). Es importante mencionar que los datos originales se encontraban en grados Kelvin, por lo que se realizó una conversión a

grados Celsius antes de su análisis posterior.

Tabla 1: Ejemplo hipotético de la manera de realizar el cálculo de la temperatura máxima y mínima el pronóstico del sistema SREF.

Día/hora	06Z - 00:00 am	09Z - 03:00 am	12Z - 06:00 am	15Z - 09:00 am	18Z - 12:00 pm	21Z - 15:00 pm	00Z - 18:00 pm	03Z 21:00 pm
01/01/23	16.8	14.4	13.2	15.4	20.4	25.2	23.1	18.4
02/01/23	17.3	15.3	12.7	15.2	19.8	24.5	23.0	17.8

En la Tabla 1, se ilustra un ejemplo de mediciones de temperatura con un intervalo de tiempo de 3 horas, tal como se obtienen del sistema SREF. Con base en las definiciones previas de la manera en la que toman lecturas los observadores (Quintas, 2000), se calculan las temperaturas máxima y mínima para el día 01/01/23. En este caso, se registra una temperatura máxima de 25.2 °C y una temperatura mínima de 12.7 °C. Este ejemplo demuestra claramente cómo se obtuvieron estos valores en el marco de la presente investigación.

En cuanto a la precipitación (mm), se utilizó el acumulado en el mismo intervalo de tiempo mencionado anteriormente.

Una vez realizado el cálculo de las variables de interés, se procedió a la extracción de los datos utilizando el software Grid Analysis and Display System (GrAds). Se utilizaron las coordenadas correspondientes a las 1570 estaciones climáticas convencionales de las cuales se disponía de datos. Utilizando estas coordenadas, se extrajo el valor correspondiente a cada punto en la malla del sistema SREF. Como resultado, se obtuvieron dos archivos para cada variable: uno con los datos observados de las estaciones climáticas convencionales y otro con los datos pronosticados extraídos del sistema SREF para las mismas estaciones consideradas.

Posteriormente, se realizó el postprocesamiento de los datos del SREF mediante la metodología propuesta, con el objetivo de corregir el sesgo de los pronósticos de las variables

meteorológicas analizadas.

Aplicación de métodos estadísticos para identificación de sesgo

En esta sección, se identificará el sesgo sistemático en los resultados de las simulaciones numéricas del sistema SREF mediante la aplicación de métodos estadísticos. El enfoque se centra en las variables de temperatura máxima, mínima y precipitación. Por lo cual, se aplicaron diversos métodos para identificar y cuantificar el sesgo y el error.

En la evaluación del sistema SREF, se calcularon diferentes métricas estadísticas que fueron definidas con anterioridad, entre las cuales se incluyen el BIAS, RMSE, MAE y el COE. Cada una de estas métricas proporciona información valiosa sobre el rendimiento del sistema en términos de la precisión y la relación con las observaciones. Con base en esto, se procedió a identificar el sesgo presente en los resultados del sistema SREF para las variables de temperatura máxima, mínima y precipitación.

Estos estadísticos, junto con el sesgo, desempeñan un papel crucial en la evaluación y cuantificación del desempeño del sistema SREF. Aunque los resultados de RMSE, MAE y COE no se presentan en esta sección, se incluyen en el Anexo 2 para proporcionar información adicional y permitir una evaluación más completa del sistema. Si bien el enfoque principal del trabajo es reducir el sesgo, es esencial considerar también estos estadísticos complementarios. Se espera que, al aplicar el Filtro de Kalman Adaptativo como postprocesamiento de los pronósticos numéricos de SREF, la magnitud del RMSE y del MAE disminuya, indicando que los valores corregidos se acercan más a las observaciones. Por otro lado, se espera que la magnitud del COE aumente ligeramente, ya que este estadístico evalúa la relación lineal entre los pronósticos y las observaciones. Estas tendencias resaltan las ventajas de aplicar el Filtro de Kalman Adaptativo para mejorar el rendimiento del sistema y obtener resultados más precisos. Por lo tanto, se recomienda revisar el Anexo 2 para obtener una visión general de los resultados del RMSE, MAE y COE.

Es importante destacar que, debido a la amplia cantidad de estaciones climáticas utilizadas en

este estudio (un total de 1570 estaciones), se optó por presentar los resultados en forma de intervalos. Esto permite una mejor interpretación de la distribución del sesgo en los resultados del sistema SREF. Los intervalos están definidos de manera consistente, abarcando valores específicos de sesgo para cada variable, permitiendo una mejor visualización y comprensión de los resultados.

Los cálculos fueron realizados por trimestres. Dichos trimestres son: diciembre, enero y febrero de 2020-2021, así como mayo, junio y julio de 2020, y se realizaron para las variables de temperatura máxima, temperatura mínima y precipitación. Esto permite analizar las diferencias estacionales en el comportamiento del sesgo.

A continuación, se presentan las tablas y figuras correspondientes a los intervalos de sesgo de temperatura máxima, temperatura mínima y precipitación, dichos productos muestran el sesgo entre el pronóstico y la observación para cada una de las variables en forma de intervalos.

Como aclaración, la fila donde contenga “[...]” representa que durante varios intervalos consecutivos se presentaron cero números de datos. Se proporciona el número de datos y el porcentaje correspondiente del total que se encuentra dentro de cada intervalo, lo que permite visualizar y comprender los resultados de manera más efectiva:

Tabla 2: Intervalos de sesgo para la temperatura máxima en el trimestre de diciembre, enero y febrero (DEF) 2020-2021, pronóstico vs observación.

INTERVALO	NUM. DATOS	%
[-20.0, -18.0]	1	0.06369427
[-18.0, -16.0]	0	0
[-16.0, -14.0]	7	0.44585987
[-14.0, -12.0]	19	1.21019108
[-12.0, -10.0]	35	2.22929936
[-10.0, -8.0]	114	7.2611465
[-8.0, -6.0]	205	13.0573248
[-6.0, -4.0]	260	16.5605096
[-4.0, -2.0]	253	16.1146497
[-2.0, 0.0]	247	15.7324841
[0.0, 2.0]	245	15.6050955
[2.0, 4.0]	101	6.43312102
[4.0, 6.0]	43	2.7388535
[6.0, 8.0]	24	1.52866242
[8.0, 10.0]	12	0.76433121
[10.0, 12.0]	2	0.12738854
[12.0, 14.0]	2	0.12738854

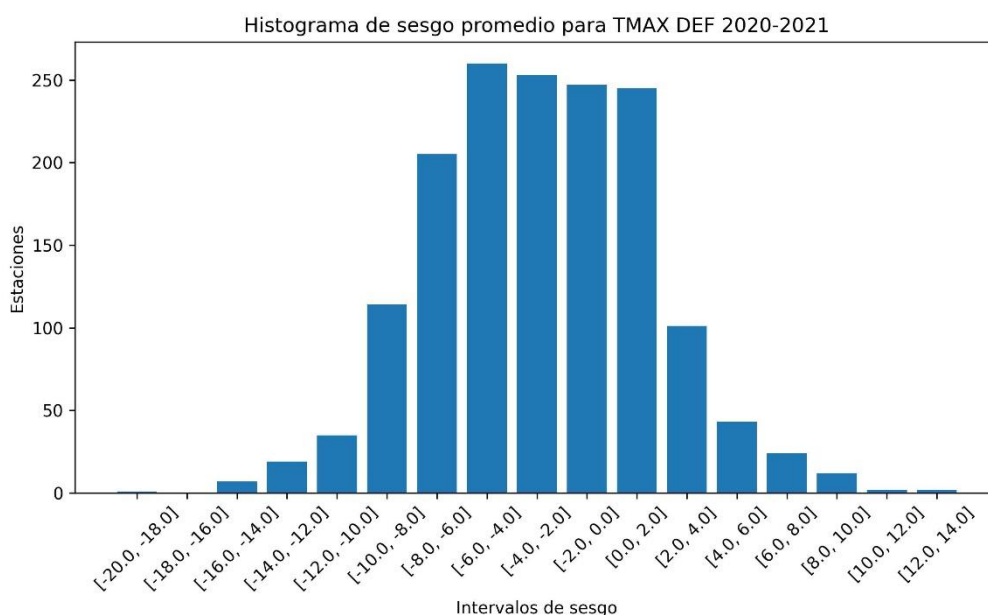


Figura 4: Histograma de sesgo promedio para la temperatura máxima en el periodo DEF 2020-2021, pronóstico vs observación.

Tabla 3: Intervalos de sesgo para la temperatura máxima en el trimestre de mayo, junio y julio 2020, pronóstico vs observación.

INTERVALO	NUM. DATOS	%
[-20.0, -18.0]	1	0.06369427
[-18.0, -16.0]	17	1.08280255
[-16.0, -14.0]	82	5.22292994
[-14.0, -12.0]	123	7.8343949
[-12.0, -10.0]	156	9.93630573
[-10.0, -8.0]	196	12.4840764
[-8.0, -6.0]	216	13.7579618
[-6.0, -4.0]	147	9.36305732
[-4.0, -2.0]	149	9.49044586
[-2.0, 0.0]	169	10.7643312
[0.0, 2.0]	212	13.5031847
[2.0, 4.0]	81	5.15923567
[4.0, 6.0]	18	1.14649682
[6.0, 8.0]	2	0.12738854
[8.0, 10.0]	1	0.06369427

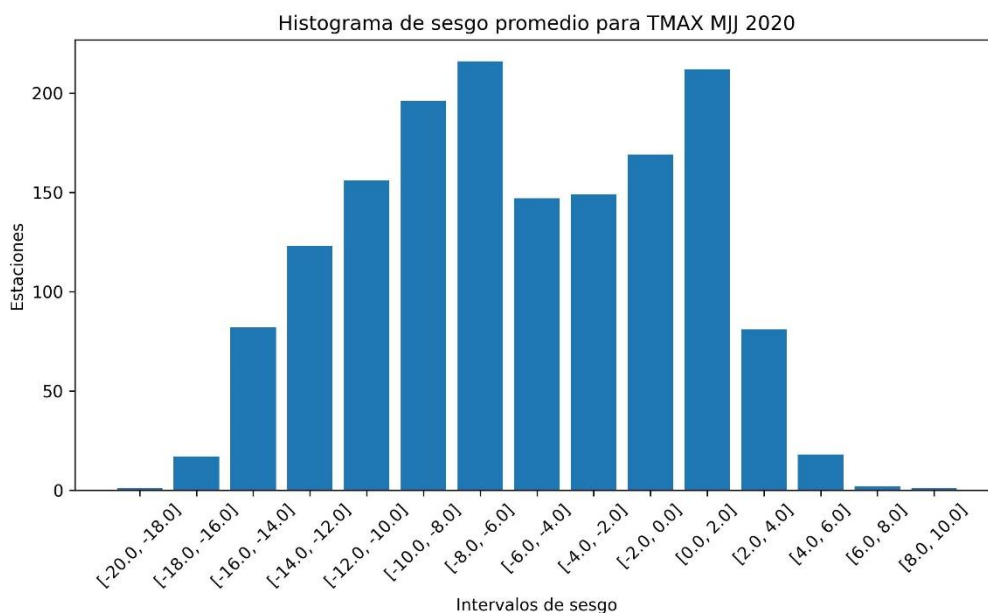


Figura 5: Histograma de sesgo promedio para la temperatura máxima en el periodo MJJ 2020, pronóstico vs observación.

Tabla 4: Intervalos de sesgo para la temperatura mínima en el trimestre de diciembre, enero y febrero 2020-2021, pronóstico vs observación.

INTERVALO	NUM. DATOS	%
[-14.0, -12.0]	2	0.12738854
[-12.0, -10.0]	11	0.70063694
[-10.0, -8.0]	37	2.3566879
[-8.0, -6.0]	22	1.40127389
[-6.0, -4.0]	22	1.40127389
[-4.0, -2.0]	61	3.88535032
[-2.0, 0.0]	236	15.0318471
[0.0, 2.0]	282	17.9617834
[2.0, 4.0]	211	13.4394904
[4.0, 6.0]	201	12.8025478
[6.0, 8.0]	158	10.0636943
[8.0, 10.0]	174	11.0828025
[10.0, 12.0]	75	4.77707006
[12.0, 14.0]	60	3.82165605
[14.0, 16.0]	12	0.76433121
[16.0, 18.0]	6	0.38216561

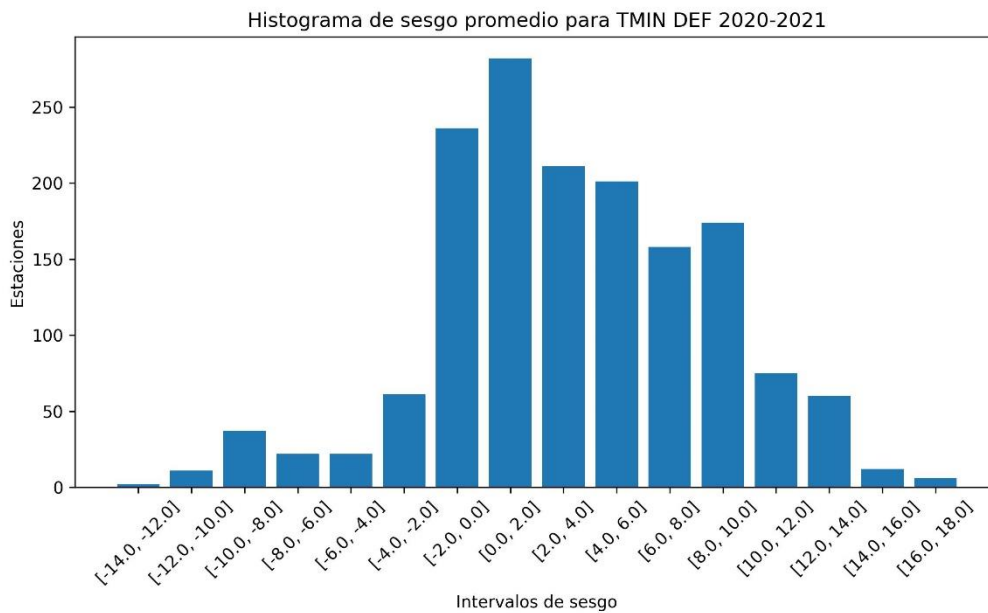


Figura 6: Histograma de sesgo promedio para la temperatura mínima en el periodo DEF 2020-2021, pronóstico vs observación.

Tabla 5: Intervalos de sesgo para la temperatura mínima en el trimestre de mayo, junio y julio 2020, pronóstico vs observación.

INTERVALO	NUM. DATOS	%
[-12.0, -10.0]	17	1.08280255
[-10.0, -8.0]	125	7.96178344
[-8.0, -6.0]	172	10.955414
[-6.0, -4.0]	195	12.4203822
[-4.0, -2.0]	228	14.522293
[-2.0, 0.0]	208	13.2484076
[0.0, 2.0]	169	10.7643312
[2.0, 4.0]	141	8.98089172
[4.0, 6.0]	207	13.1847134
[6.0, 8.0]	88	5.60509554
[8.0, 10.0]	18	1.14649682
[10.0, 12.0]	2	0.12738854

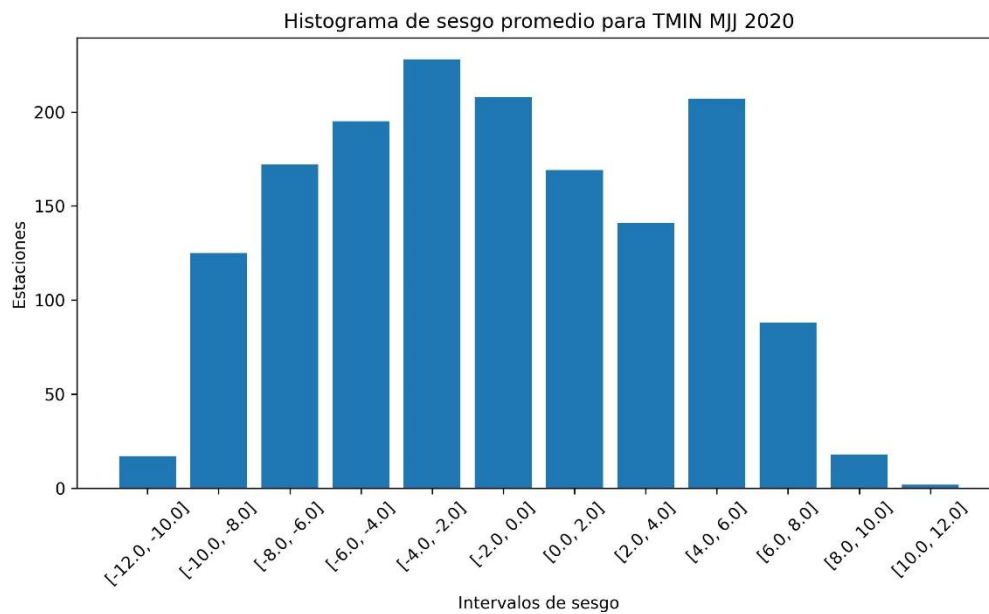


Figura 7: Histograma de sesgo promedio para la temperatura mínima en el periodo MJJ 2020, pronóstico vs observación.

Tabla 6: Intervalos de sesgo para la precipitación en el trimestre de diciembre, enero y febrero 2020-2021, pronóstico vs observación.

INTERVALO	NUM. DATOS	%
[-2.0, 0.0]	29	1.84713376
[0.0, 2.0]	1214	77.3248408
[2.0, 4.0]	208	13.2484076
[4.0, 6.0]	58	3.69426752
[6.0, 8.0]	40	2.5477707
[8.0, 10.0]	14	0.89171975
[10.0, 12.0]	4	0.25477707
[12.0, 14.0]	3	0.1910828

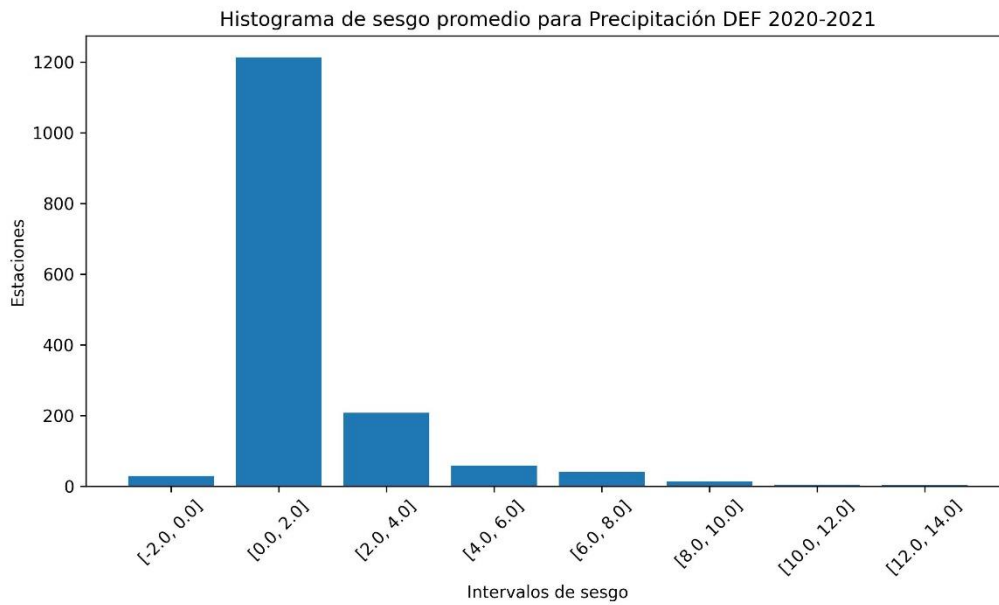


Figura 8: Histograma de sesgo promedio para la precipitación en el periodo DEF 2020-2021, pronóstico vs observación.

Tabla 7: Intervalos de sesgo para la precipitación en el trimestre de mayo, junio y julio 2020, pronóstico vs observación.

INTERVALO	NUM. DATOS	%
[-4.0, -2.0]	22	1.40127389
[-2.0, 0.0]	268	17.0700637
[0.0, 2.0]	237	15.0955414
[2.0, 4.0]	280	17.8343949
[4.0, 6.0]	196	12.4840764
[6.0, 8.0]	129	8.21656051
[8.0, 10.0]	122	7.77070064
[10.0, 12.0]	67	4.26751592
[12.0, 14.0]	55	3.50318471
[14.0, 16.0]	53	3.37579618
[16.0, 18.0]	33	2.10191083
[18.0, 20.0]	25	1.59235669
[20.0, 22.0]	17	1.08280255
[22.0, 24.0]	24	1.52866242
[24.0, 26.0]	15	0.95541401
[26.0, 28.0]	3	0.1910828
[28.0, 30.0]	4	0.25477707
[30.0, 32.0]	3	0.1910828
[32.0, 34.0]	0	0
[34.0, 36.0]	2	0.12738854
[36.0, 38.0]	4	0.25477707
[38.0, 40.0]	0	0
[40.0, 42.0]	2	0.12738854
[42.0, 44.0]	3	0.1910828
...
[50.0, 52.0]	4	0.25477707
[52.0, 54.0]	2	0.12738854

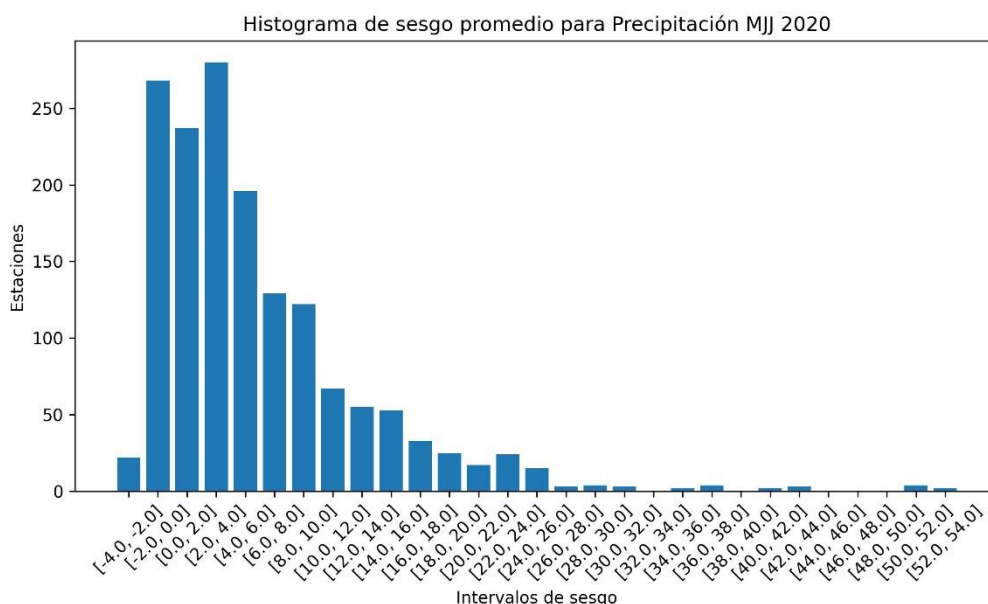


Figura 9: Histograma de sesgo promedio para la precipitación en el periodo MJJ 2020, pronóstico vs observación.

En la Tabla 2 – Figura 4, se presentan los resultados de los intervalos de sesgo para la temperatura máxima (DEF). Se identificaron intervalos que van desde -20 a 14 °C. El mayor porcentaje de los datos se encuentran en el rango de intervalo de [-10 a 4.0]. En particular, se observa que el 16.56% de los datos presentan un sesgo entre [-6.0 a -4.0], lo cual representa 260 estaciones de las 1570 estaciones analizadas.

En la Tabla 3 – Figura 5, se presentan los intervalos de sesgo para la temperatura máxima (MJJ), con un rango de -20 a 10 °C. El mayor porcentaje de los datos se encuentran en el intervalo de [-14 a 2.0]. Es notable que el 13.75% de los datos presentan un sesgo entre [-8.0 a -6.0], lo cual corresponde a 216 estaciones de las 1570 estaciones analizadas.

La Tabla 4 – Figura 6, muestra los intervalos de sesgo para la temperatura mínima (DEF), que van desde -14 a 18 °C. La mayor concentración de datos se encuentra en el intervalo de [-2 a 10]. Se destaca que el 17.96% de los datos presentan un sesgo entre [0.0 a 2.0], lo cual

representa 282 estaciones de las 1570 estaciones analizadas.

En la Tabla 5 – Figura 7, se presentan los intervalos de sesgo para la temperatura mínima (MJJ), abarcando un rango de -12 a 12 °C. La mayor cantidad de los datos se encuentran en el intervalo de [-10 a 6.0]. Es relevante mencionar que el 14.52% de los datos presentan un sesgo entre [-4.0 a -2.0], lo cual equivale a 228 estaciones de las 1570 estaciones analizadas.

La Tabla 6 – Figura 8, muestra los intervalos de sesgo para la precipitación (DEF), que varían de -2 a 14 mm. El porcentaje mayor de datos se encuentran en el intervalo de [0 a 4]. Es notable que el 77.32% de los datos presenten un sesgo entre [0.0 a 2.0], lo cual representa 1214 estaciones de las 1570 estaciones analizadas.

En la Tabla 7 – Figura 9, se presentan los intervalos de sesgo para la precipitación (MJJ), con un rango de -4 a 54 mm. La mayor concentración de datos se encuentra en el intervalo de [-2 a 10]. Es interesante observar que el 17.83% de los datos presentan un sesgo entre [2.0 a 4.0], lo cual corresponde a 280 estaciones de las 1570 estaciones analizadas.

Estos resultados muestran los intervalos de sesgo para diferentes variables analizadas en diferentes períodos de tiempo. Se identificaron los rangos de valores en los que se concentran la mayoría de los datos para la temperatura máxima, temperatura mínima y precipitación. Además, se resalta el porcentaje de datos que presentan un sesgo dentro de intervalos específicos. Estos hallazgos son basados en el análisis de 1570 estaciones y proporcionan información importante sobre la precisión de los pronósticos en relación con las observaciones.

Tabla 8: Valor de sesgo máximo positivo y negativo encontrado para cada caso de estudio, pronóstico vs observación.

Variable - periodo	Máximo sesgo positivo encontrado	Máximo sesgo negativo encontrado
Temperatura máxima - DEF	12.7	-19.75
Temperatura máxima - MJJ	8.37	-18.28
Temperatura mínima - DEF	17.77	-13.49
Temperatura mínima - MJJ	10.7	-10.54
Precipitación - DEF	13.36	-0.61
Precipitación - MJJ	52.57	-3.3

Como se puede observar en el conjunto de tablas (2 a 7), los intervalos de sesgo abarcan un amplio rango de valores, en la tabla 8 podemos observar el sesgo máximo y mínimo identificado en cada caso de estudio. Esto indica que el sistema SREF presenta un sesgo significativamente alto en sus pronósticos. A continuación, en la Tabla 9, se presentan los promedios de los intervalos de sesgo, además, del sesgo promedio y el sesgo absoluto promedio para cada variable:

Tabla 9: Resumen de sesgos e intervalos para las variables de interés

Variable	Intervalo promedio	Sesgo promedio	Sesgo absoluto promedio
Temperatura	[-16.5 a 13.5]	-1.59 °C	4.92 °C
Temperatura máxima	[-20 a 12]	-4.25 °C	5.27 °C
Temperatura mínima	[-13 a 14]	1.07 °C	4.35 °C
Temperatura (DEF)	[-17 a 16]	0.28 °C	4.48 °C
Temperatura (MJJ)	[-16 a 11.0]	-3.45 °C	5.35 °C
Precipitación	[-3 a 34]	3.66 mm	3.87 mm
Precipitación (DEF)	-	1.40 mm	1.40 mm
Precipitación (MJJ)	-	5.92 mm	6.33 mm

Es importante destacar que el cálculo del sesgo se realizó considerando el trimestre completo de información. En el caso de la precipitación, se observó la presencia de numerosos días con precipitaciones escasas o nulas, lo cual puede influir en la baja magnitud del sesgo identificado. Sin embargo, como complemento, se decidió calcular el sesgo únicamente considerando los días con precipitaciones mayores a 0.1 mm, siguiendo el enfoque utilizado

por Velásquez et al. (2020), donde no se consideran las precipitaciones nulas en el análisis.

Al realizar este enfoque, se obtuvieron resultados modificados para el sesgo promedio de los meses de MJJ. El sesgo promedio aumentó de 5.92 mm a 6.24 mm, mientras que el sesgo absoluto pasó de 6.33 mm a 6.79 mm. Para los meses de DEF, se observó un cambio similar en los resultados. El sesgo promedio pasó de 1.40 mm a 2.58 mm, y el sesgo absoluto aumentó de 1.40 mm a 2.74 mm.

Estos amplios intervalos de sesgo resaltan la necesidad de aplicar un método de corrección de sesgo para mejorar la precisión de los pronósticos.

Tabla 10 : Distribución del sesgo positivo y negativo para cada una de las variables y periodos analizados, identificación del comportamiento del sistema SREF.

Variable - periodo	Número de datos con sesgo < 0	Número de datos con sesgo > 0	Porcentaje de datos con sesgo < 0	Porcentaje de datos con sesgo > 0	Comportamiento del modelo
Temperatura máxima - DEF	1141	429	72.68	27.32	Subestima
Temperatura máxima - MJJ	1256	314	80.00	20.00	Subestima
Temperatura mínima - DEF	391	1179	24.90	75.10	Sobrestima
Temperatura mínima - MJJ	945	625	60.19	39.81	Subestima
Precipitación - DEF	29	1541	1.85	98.15	Sobrestima
Precipitación - MJJ	290	1280	18.47	81.53	Sobrestima

Retomando la definición previamente establecida de BIAS (sesgo), que se refiere a la tendencia de un modelo a sobrestimar o subestimar una variable, la tabla 10 presenta de manera detallada la distribución de los sesgos y las tendencias observadas en el sistema SREF para los casos analizados. La tabla muestra la cantidad de datos con sesgos menores a cero y mayores a cero para cada variable y periodo, y también indica el porcentaje correspondiente de cada grupo en relación al total de datos. Además, se identifica la tendencia del modelo (subestima o sobrestima) para cada combinación de variable y periodo. Esta tabla brinda una visión precisa del comportamiento del sesgo en el sistema SREF, lo que ayuda a identificar su tendencia en estos casos específicos de estudio.

De acuerdo con los datos presentados en la tabla 10, se puede apreciar una tendencia del sistema SREF a subestimar la temperatura máxima en ambos trimestres y la temperatura mínima durante los meses de MJJ. Por otro lado, se observa que el sistema tiende a sobreestimar la temperatura mínima durante el periodo DEF, así como la precipitación durante los periodos DEF y MJJ.

Para aclarar estos conceptos, consideremos el siguiente ejemplo: supongamos que el sistema SREF pronostica una temperatura mínima de 12.5 °C para un día de invierno y una ubicación específica. Al decir que el sistema SREF está sobreestimando la temperatura mínima, nos referimos a que está dando valores más altos en comparación con lo que reportarían las observaciones. En este caso hipotético, imaginemos que se registró una temperatura mínima de 10.4 °C. Es evidente que la temperatura pronosticada de 12.5 °C por el SREF es superior al valor observado. Por lo tanto, concluimos que el sistema SREF sobreestimó la temperatura mínima para ese día en particular.

Aplicación del método de corrección de sesgo

En esta sección, se llevará a cabo la aplicación del método de corrección de sesgo utilizando el Filtro de Kalman Adaptativo, como se describió detalladamente en la sección “Método de corrección basado en filtros de Kalman”, las ecuaciones se resumen en la Figura 10. El objetivo principal es corregir los pronósticos del modelo SREF para las variables de temperatura máxima, mínima y precipitación.

Las ecuaciones y algoritmos descritos en la sección 2.2 se implementaron en un código en Python (fig. 11). Se emplearon nuevamente los trimestres previamente mencionados. Para realizar diversos experimentos considerando estos casos de estudio. Reafirmando que los períodos de tiempo considerados fueron: diciembre, enero y febrero de 2020-2021, y mayo, junio y julio de 2020 para las variables de temperatura máxima, mínima y precipitación.

Para cada caso de estudio trimestral, se ejecutó el código correspondiente (ver fig. 11) y como resultado, se obtuvo un archivo CSV que contiene información detallada para cada una de las estaciones climatológicas convencionales (fig. 12). Los datos recopilados incluyen el nombre del municipio, el estado, la latitud, la longitud, la fecha que representa el día, la temperatura pronosticada, la temperatura observada y la temperatura corregida mediante el Filtro de Kalman Adaptativo. Para una mejor comprensión del archivo, se recomienda consultar la Figura 12.

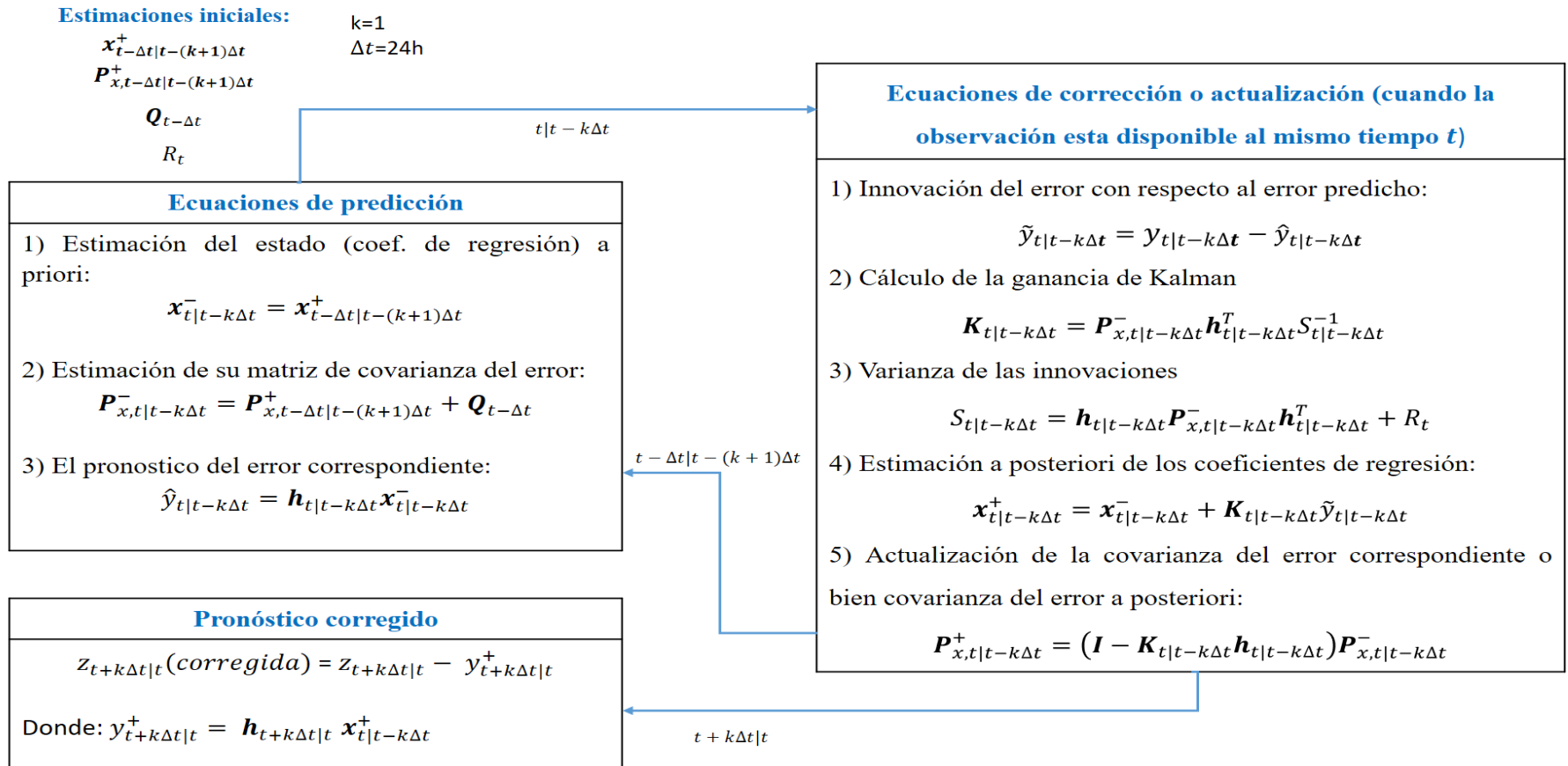


Figura 10: Ecuaciones del Filtro de Kalman Adaptativo utilizadas para realizar la corrección de sesgo.

Al analizar las ecuaciones proporcionadas en la Figura 10, es importante destacar que las ecuaciones de actualización o corrección tienen un impacto a partir del segundo día. Por lo tanto, es comprensible que la corrección para el primer día de análisis sea deficiente, ya que los coeficientes de correlación aún no han sido actualizados y no han influido en el desarrollo del Filtro. En la Figura 12, se puede observar que la temperatura corregida para el día 1 es muy distante de la observación registrada, con un valor de 8.65 °C en comparación con los 23.2 °C observados. Sin embargo, a partir del segundo día, una vez que se han actualizado los coeficientes, la corrección mejora significativamente, lo que respalda la afirmación de Galanis & Anadranistokis, 2002, de que los coeficientes de correlación tienden a acercarse rápidamente a su valor "real".

```

# Predicción
x = x
P = P + Q_lista[dia-1]
Yg = np.vdot(h, x)
Yg_lista.append(Yg)
Tcorre=sref-Yg

# Corrección o actualización
e=sref-obs
e_lista.append(e)
#Yo = (srefant-obsant) - Yg
Yo = (sref-obs) - Yg
Yo_lista.append(Yo)
S = np.vdot(np.dot(h,P),np.transpose(h)) + R_lista[dia-1]
K = np.dot(P,np.transpose(h))*1/S
x = x + (K*Yo)
I = np.identity(2)
Kh = np.dot(K,h)
IKh = I-Kh
P = np.dot(IKh, P)
P_lista.append(P)
Ypos = np.dot(h, x)
#Tcorre = sref-(Ypos[0][0])
x_lista.append(x.astype(np.float64))

# Cálculo de Q_lista
if dia < 3:
    Q_lista.append(Q)
elif dia < 7:
    indices = range(dia-3, dia)
    Q_t = np.sum(np.subtract([x_lista[i] for i in indices], [x_lista[i-1] for i in indices]), axis=0) / 3
    Q_t = np.sum(np.subtract([x_lista[i] for i in indices], [x_lista[i-1] for i in indices] - Q_t)**2, axis=0) / 3
    Q_t = 1/3*Q_t
    Q_t = I*Q_t
    Q_lista.append(Q_t)
else:
    indices = range(dia-7, dia)
    Q_t = np.sum(np.subtract([x_lista[i] for i in indices], [x_lista[i-1] for i in indices]), axis=0) / 7
    Q_t = np.sum(np.subtract([x_lista[i] for i in indices], [x_lista[i-1] for i in indices] - Q_t)**2, axis=0) / 7
    Q_t = 1/6*Q_t
    Q_t = I*Q_t
    Q_lista.append(Q_t)

```

Figura 11: Algoritmo de Filtro de Kalman Adaptativo escrito en Python.

	A	B	C	D	E	F	G	H	
1	mpio	edo	lat	lon	fecha	obs	sref	corregido	
2	Aguascaliente	Ags.	21.739	-102.184	1	23.2	19.7	8.65	
3	Aguascaliente	Ags.	21.739	-102.184	2	23.7	19.8	23.29	
4	Aguascaliente	Ags.	21.739	-102.184	3	19.9	19.8	23.7	
5	Aguascaliente	Ags.	21.739	-102.184	4	18.7	15.2	15.03	
6	Aguascaliente	Ags.	21.739	-102.184	5	19.6	18.5	21	
7	Aguascaliente	Ags.	21.739	-102.184	6	20	19.08	20.33	
8	Aguascaliente	Ags.	21.739	-102.184	7	21.1	19.31	20.38	
9	Aguascaliente	Ags.	21.739	-102.184	8	23	19.3	20.92	
10	Aguascaliente	Ags.	21.739	-102.184	9	25.2	20.4	23.63	
11	Aguascaliente	Ags.	21.739	-102.184	10	25.7	21.3	26.25	
12	Aguascaliente	Ags.	21.739	-102.184	11	25.6	21.4	26.02	
13	Aguascaliente	Ags.	21.739	-102.184	12	25.5	21.4	25.71	
14	Aguascaliente	Ags.	21.739	-102.184	13	25.8	22	26.29	
15	Aguascaliente	Ags.	21.739	-102.184	14	22.7	21.2	24.96	
16	Aguascaliente	Ags.	21.739	-102.184	15	24.4	21	23.25	
17	Aguascaliente	Ags.	21.739	-102.184	16	21.9	20.2	22.86	
18	Aguascaliente	Ags.	21.739	-102.184	17	22.4	20.5	22.66	
19	Aguascaliente	Ags.	21.739	-102.184	18	24.4	20.8	22.86	
20	Aguascaliente	Ags.	21.739	-102.184	19	25	20.8	23.71	
21	Aguascaliente	Ags.	21.739	-102.184	20	24.1	21.5	25.13	
22	Aguascaliente	Ags.	21.739	-102.184	21	24.6	21.9	25.2	
23	Aguascaliente	Ags.	21.739	-102.184	22	25.2	21.8	24.82	
24	Aguascaliente	Ags.	21.739	-102.184	23	26.7	22.5	25.89	
25	Aguascaliente	Ags.	21.739	-102.184	24	21	21.1	24.56	
26	Aguascaliente	Ags.	21.739	-102.184	25	21.4	20	21.17	
27	Aguascaliente	Ags.	21.739	-102.184	26	23.6	21.3	23.01	
28	Aguascaliente	Ags.	21.739	-102.184	27	25.6	20.8	22.76	

Figura 12: Ejemplo de los resultados obtenidos al aplicar el código en Python.

Recordando lo mencionado en Galanis & Anadranistokis (2002), es importante destacar que los coeficientes de regresión y la covarianza de su error (\mathbf{x} y \mathbf{P}) tienden a converger rápidamente hacia sus valores verdaderos en las primeras actualizaciones. Por lo que, para nuestro caso, como \mathbf{P} y \mathbf{x} , se tomaron unos coeficientes arbitrarios, debido a la suposición de que convergerán a su valor real en actualizaciones próximas.

Sin embargo, es crucial tener cuidado al seleccionar las covarianzas de los errores de las ecuaciones de estado y observación (\mathbf{Q} y \mathbf{R}), ya que tienen un impacto significativo en el desempeño general del filtro. Por tanto, es fundamental una adecuada inicialización de estos parámetros para lograr una estimación precisa para las variables corregidas. Para esta investigación, se tomó lo propuesto por Galanis & Anadranistokis (2002), como fue mencionado en la sección 2.2, las ecuaciones (18) y (19) permiten realizar el cálculo de estos parámetros, basándose en la información de los últimos 7 días, como lo son los coeficientes de regresión predichos (\mathbf{Q}) y las innovaciones de los errores de pronóstico (\mathbf{R}).

Con base en lo anterior, para calcular \mathbf{Q} y \mathbf{R} , se siguió un procedimiento basado en días

previos. Inicialmente, para los primeros tres días, se utilizó una matriz identidad $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ como aproximación para la covarianza del error de la ecuación de estado (\mathbf{Q}). En cuanto a la covarianza del error de la ecuación de observación (\mathbf{R}), se utilizó el RMSE entre el pronóstico y la observación correspondiente al mes de inicio del análisis. Por ejemplo, para el trimestre DEF, se utilizó para los primeros 3 días el RMSE calculado para el mes de diciembre para cada una de las estaciones.

Posteriormente, a medida que se disponía de información de tres días anteriores, se procedió a calcular estas covarianzas de errores utilizando las ecuaciones 18 y 19, siguiendo el enfoque utilizado en el estudio de Stathopoulos et al. (2013), donde fueron aplicadas haciendo uso de 3 días previos de información. Se continuó iterando este proceso en los días sucesivos hasta contar con un total de siete días de cálculos, momento en el cual se aplicaron las ecuaciones según se proponen.

La Figura 13 brinda una representación visual del procedimiento empleado para el cálculo de \mathbf{Q} y \mathbf{R} . En los días 6, 7 y 8 (color azul), se utilizó la matriz identidad para \mathbf{Q} , mientras que para la \mathbf{R} se empleó el RMSE. A partir del día 9 hasta el día 12, se aplicaron las ecuaciones considerando un período de tres días de información. Finalmente, a partir del día 13 en adelante (color rojo), se aplicaron las ecuaciones 18 y 19 tal como se presentan en el estudio.

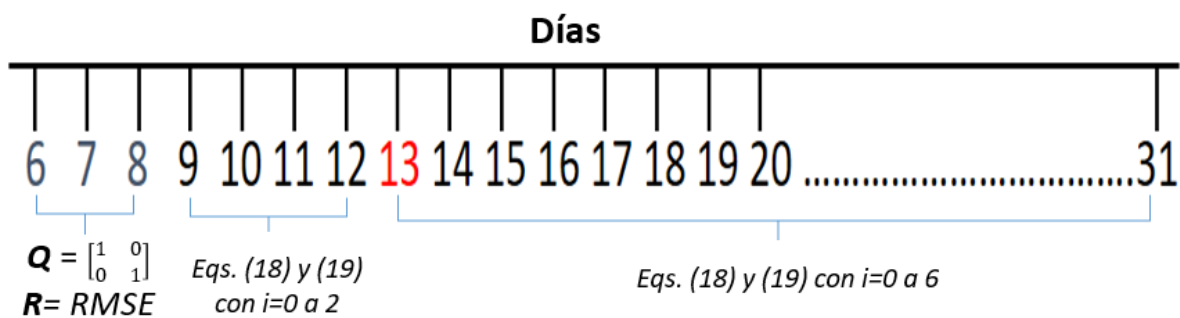


Figura 13: Esquematización de la manera en la que se calculó la covarianza del error de la ecuación de estado $\mathbf{Q}_{t-k\Delta t}$ y la covarianza del error de la ecuación de observación \mathbf{R}_t , según lo propuesto por (Galanis & Anadranistokis , 2002).

A continuación, se presenta un ejemplo para ilustrar el proceso de corrección de sesgo en un

día específico, proporcionando un mayor entendimiento de la aplicación práctica de la metodología para la corrección de las variables meteorológicas analizadas.

Se calcularon las covarianzas del error de la ecuación de estado y de observación (Q y R), utilizando las ecuaciones 18 y 19, respectivamente. Para este cálculo se empleó información de los 7 días previos. Debido al extenso tamaño de las ecuaciones, no se incluyó el desarrollo detallado en el ejemplo de a continuación, sin embargo, fueron aplicadas correctamente en el análisis.

Datos:

Estación: Zacatecas, Zac. $k=1$, $\Delta t=24h$

Estimaciones iniciales:

$$\begin{array}{l} \mathbf{x}_{9|8}^+ \\ \mathbf{P}_{x,9|8}^+ \\ \mathbf{Q}_9 \\ R_{10} \end{array} \quad \begin{array}{l} \begin{bmatrix} 1.22 \\ -0.56 \end{bmatrix} \\ \begin{bmatrix} 6.93 & -0.32 \\ -0.32 & 0.01 \end{bmatrix} \\ \begin{bmatrix} 0.0003 & 0 \\ 0 & 0.0009 \end{bmatrix} \\ 0.2159 \end{array}$$

Temperatura máxima pronosticada para el día 10 de mayo del 2020, emitida el día 9 ($\mathbf{z}_{10|9}$):

23.1 °C

Temperatura máxima pronosticada para el día 11 de mayo del 2020, emitida el día 10

($\mathbf{z}_{11|10}$): **23.4 °C**

Temperatura máxima observada para el día 10 de mayo del 2020, (\mathbf{Obs}_{10}): **31.9 °C**

Temperatura máxima observada para el día 11 de mayo del 2020, (\mathbf{Obs}_{11}): **32.0 °C**

Paso de predicción:

1. Estimación del estado a priori:

$$\mathbf{x}_{t|t-k\Delta t}^- = \mathbf{x}_{t-\Delta t|t-(k+1)\Delta t}^+$$

$$\mathbf{x}_{10|9}^- = \mathbf{x}_{9|8}^+$$

$$\mathbf{x}_{10|9}^- = \begin{bmatrix} 1.22 \\ -0.56 \end{bmatrix}$$

2. Estimación de su matriz de covarianza del error:

$$\begin{aligned} \mathbf{P}_{x,t|t-k\Delta t}^- &= \mathbf{P}_{x,t-\Delta t|t-(k+1)\Delta t}^+ + \mathbf{Q}_{t-\Delta t} \\ \mathbf{P}_{x,10|9}^- &= \mathbf{P}_{x,9|8}^+ + \mathbf{Q}_9 \\ \mathbf{P}_{x,10|9}^- &= \begin{bmatrix} 6.93 & -0.32 \\ -0.32 & 0.01 \end{bmatrix} + \begin{bmatrix} 0.0003 & 0 \\ 0 & 0.0009 \end{bmatrix} \\ \mathbf{P}_{x,10|9}^- &= \begin{bmatrix} 6.9303 & -0.32 \\ -0.32 & 0.0109 \end{bmatrix} \end{aligned}$$

3. El pronóstico del error correspondiente:

$$\begin{aligned} \hat{y}_{t|t-k\Delta t} &= \mathbf{h}_{t|t-k\Delta t} \mathbf{x}_{t|t-k\Delta t}^- \\ \hat{y}_{10|9} &= \mathbf{h}_{10|9} \mathbf{x}_{10|9}^- \\ \hat{y}_{10|9} &= [1 \quad 23.1 \text{ °C}] * \begin{bmatrix} 1.22 \\ -0.56 \end{bmatrix} \\ \hat{y}_{10|9} &= -11.71 \text{ °C} \end{aligned}$$

Una vez concluye el día, y tenemos la observación disponible para el tiempo t. Las ecuaciones de actualización se calculan de la siguiente manera:

1. Innovación del error con respecto al error predicho:

$$\begin{aligned} \tilde{y}_{t|t-k\Delta t} &= z_{t|t-k\Delta t} - Obs_t - \hat{y}_{t|t-k\Delta t} \\ \tilde{y}_{10|9} &= z_{10|9} - Obs_{10} - \hat{y}_{10|9} \\ \tilde{y}_{10|9} &= (23.1\text{°C} - 31.9 \text{ °C}) - (-11.71 \text{ °C}) \\ \tilde{y}_{10|9} &= 2.91 \text{ °C} \end{aligned}$$

2. Varianza de las innovaciones:

$$S_{t|t-k\Delta t} = \mathbf{h}_{t|t-k\Delta t} \mathbf{P}_{x,t|t-k\Delta t}^- \mathbf{h}_{t|t-k\Delta t}^T + R_t$$

$$S_{10|9} = \mathbf{h}_{10|9} \mathbf{P}_{x,10|9}^- \mathbf{h}_{10|9}^T + R_{10}$$

$$S_{10|9} = [1 \quad 23.1 \text{ }^\circ\text{C}] * \begin{bmatrix} 6.9303 & -0.32 \\ -0.32 & 0.0109 \end{bmatrix} * \begin{bmatrix} 1 \\ 23.1 \text{ }^\circ\text{C} \end{bmatrix} + 0.2159$$

$$S_{10|9} = -1.82$$

3. Cálculo de la ganancia de Kalman:

$$\mathbf{K}_{t|t-k\Delta t} = \mathbf{P}_{x,t|t-k\Delta t}^- \mathbf{h}_{t|t-k\Delta t}^T \mathbf{S}_{t|t-k\Delta t}^{-1}$$

$$\mathbf{K}_{10|9} = \mathbf{P}_{x,10|9}^- \mathbf{h}_{10|9}^T \mathbf{S}_{10|9}^{-1}$$

$$\mathbf{K}_{10|9} = \begin{bmatrix} 6.9303 & -0.32 \\ -0.32 & 0.0109 \end{bmatrix} * \begin{bmatrix} 1 \\ 23.1 \text{ }^\circ\text{C} \end{bmatrix} * -1.82^{-1}$$

$$\mathbf{K}_{10|9} = \begin{bmatrix} 0.253 \\ 0.037 \end{bmatrix}$$

4. Estimación a posteriori de los coeficientes de regresión:

$$\mathbf{x}_{t|t-k\Delta t}^+ = \mathbf{x}_{t|t-k\Delta t}^- + \mathbf{K}_{t|t-k\Delta t} \tilde{\mathbf{y}}_{t|t-k\Delta t}$$

$$\mathbf{x}_{10|9}^+ = \mathbf{x}_{10|9}^- + \mathbf{K}_{10|9} \tilde{\mathbf{y}}_{10|9}$$

$$\mathbf{x}_{10|9}^+ = \begin{bmatrix} 1.22 \\ -0.56 \end{bmatrix} + \begin{bmatrix} 0.253 \\ 0.037 \end{bmatrix} * 2.91 \text{ }^\circ\text{C}$$

$$\mathbf{x}_{10|9}^+ = \begin{bmatrix} 1.95 \\ -0.45 \end{bmatrix}$$

5) Actualización de la covarianza del error correspondiente o bien covarianza del error a posteriori:

$$\mathbf{P}_{x,t|t-k\Delta t}^+ = (\mathbf{I} - \mathbf{K}_{t|t-k\Delta t} \mathbf{h}_{t|t-k\Delta t}) \mathbf{P}_{x,t|t-k\Delta t}^-$$

$$\mathbf{P}_{x,10|9}^+ = (\mathbf{I} - \mathbf{K}_{10|9} \mathbf{h}_{10|9}) \mathbf{P}_{x,10|9}^-$$

$$\mathbf{P}_{x,10|9}^+ = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.253 \\ 0.037 \end{bmatrix} * [1 \quad 23.1 \text{ }^\circ\text{C}] \right) * \begin{bmatrix} 6.9303 & -0.32 \\ -0.32 & 0.0109 \end{bmatrix}$$

$$\mathbf{P}_{x,10|9}^+ = \begin{bmatrix} 7.04 & -0.30 \\ -0.30 & 0.013 \end{bmatrix}$$

Corrección del pronóstico de temperatura máxima para el día 11 de mayo para la estación de

Zacatecas, Zac., en $t + k\Delta t|t$, utilizando los coeficientes de correlación actualizados en el tiempo $t|t - k\Delta t$, $\mathbf{x}_{t|t-k\Delta t}^+$:

$$\begin{aligned}
 z_{t+k\Delta t|t}(\text{corregida}) &= z_{t+k\Delta t|t} - \mathbf{h}_{t+k\Delta t|t} \mathbf{x}_{t|t-k\Delta t}^+ \\
 z_{11|10}(\text{corregida}) &= z_{11|10} - \mathbf{h}_{11|10} \mathbf{x}_{10|9}^+ \\
 z_{11|10}(\text{corregida}) &= 23.4 \text{ }^\circ\text{C} - ([1 \quad 23.4 \text{ }^\circ\text{C}] * \begin{bmatrix} 1.95 \\ -0.45 \end{bmatrix}) \\
 z_{11|10}(\text{corregida}) &= 23.4 \text{ }^\circ\text{C} - (-8.58) \\
 z_{11|10}(\text{corregida}) &= \mathbf{31.98 \text{ }^\circ\text{C}}
 \end{aligned}$$

Por lo tanto, el pronóstico de temperatura máxima corregida a través de la metodología de Filtro de Kalman Adaptativo, para la estación de Zacatecas el día 11 de mayo del 2020, es de 31.98 °C. Es notable mencionar, que el pronóstico proporcionado por SREF para esta variable fue de 23.4 °. Que al compararlo con la observación registrada ese día de 32°C, nos da un error de pronóstico de -8.6 °C. Pero si se realiza la misma comparación con el pronóstico corregido, encontramos un error de pronóstico de -0.02 °C. Esta mejora es muy importante y destaca la eficacia del filtro utilizado. Además, es relevante mencionar que este método puede ser implementado en tiempo real para corregir pronósticos, siempre y cuando se cuenten con observaciones actualizadas diariamente, lo que garantizará un rendimiento óptimo del filtro.

Los coeficientes actualizados de $\mathbf{x}_{10|9}^+$ y $\mathbf{P}_{x,10|9}^+$, pasarían a ser las estimaciones a priori utilizadas en las ecuaciones de predicción para el siguiente día de la siguiente manera:

$$\begin{aligned}
 \mathbf{x}_{11|10}^- &= \mathbf{x}_{10|9}^+ \\
 \mathbf{x}_{10|9}^- &= \begin{bmatrix} 1.95 \\ -0.45 \end{bmatrix} \\
 \mathbf{P}_{x,11|10}^- &= \mathbf{P}_{x,10|9}^+ + \mathbf{Q}_{10} \\
 \mathbf{P}_{x,10|9}^- &= \begin{bmatrix} 7.04 & -0.30 \\ -0.30 & 0.013 \end{bmatrix} + \mathbf{Q}_{10}
 \end{aligned}$$

Se ha descrito la aplicación del Filtro de Kalman Adaptativo como un método de

postprocesamiento para corregir resultados de simulaciones numéricas del tiempo. Los pasos descritos en esta sección se repiten en un proceso iterativo. En cada iteración, se calcula el pronóstico corregido utilizando los coeficientes de correlación actualizados, y los parámetros se utilizan como estimaciones a posteriori en el siguiente paso de tiempo, previo a la obtención de la observación. Una vez registrada la observación, se procede a actualizar nuevamente los coeficientes para corregir el pronóstico en el siguiente paso de tiempo. De esta manera, se establece un ciclo continuo de corrección y actualización, que se basa en la disponibilidad de datos, con el objetivo de mejorar la precisión del pronóstico numérico a medida que se van obteniendo nuevas observaciones.

Durante el desarrollo de esta investigación, se han tenido en cuenta varias consideraciones importantes, especialmente en relación a la variable de precipitación. Al implementar el filtro, se observó que se obtenían valores de precipitación menores a cero, lo cual contradice el sentido físico de dicha variable. Esto ocurre debido a que el filtro de Kalman es un método matemático que busca obtener la mejor estimación posible basada en el pronóstico y las observaciones, sin tomar en cuenta directamente las características físicas de la variable.

Esta problemática ha sido abordada por otros investigadores, como Neslihanoglu et al., (2021), quienes también encontraron que el uso del filtro de Kalman para estimar la precipitación resultaba en valores negativos o muy cercanos a cero. Aunque intentaron solucionar el problema mediante la transformación logarítmica, no lograron resolverlo por completo. Como resultado, tomaron la decisión de considerar esos valores negativos como cero, indicando la ausencia de expectativas de precipitación para esos días.

En nuestro caso, hemos adoptado la misma medida o consideración que Neslihanoglu et al., (2021), al enfrentar esta problemática. Optamos por considerar los valores corregidos negativos de precipitación como cero, siguiendo su enfoque, con el objetivo de mantener la coherencia con las características físicas de la variable y concluir que cualquier valor negativo de precipitación indica la ausencia de probabilidad de lluvia.

Capítulo 4. Análisis de resultados

En el siguiente capítulo, se analizan los resultados obtenidos mediante la implementación de esta metodología a las variables y periodos mencionados. En este análisis, evaluaremos la mejora lograda al aplicar el Filtro de Kalman Adaptativo al comparar los pronósticos corregidos con las observaciones registradas. Mediante este análisis, es posible determinar la reducción de sesgo al utilizar el método de corrección, en comparación con las salidas en bruto del sistema SREF. Esto nos permitirá obtener una visión clara de la eficacia de la metodología implementada y su impacto en la reducción de sesgo de los pronósticos, proporcionando resultados sólidos para respaldar nuestras conclusiones.

Se compararon los valores pronosticados con los valores observados, calculando el sesgo y agrupando los resultados en intervalos de colores, rojo para sesgos positivos, blanco para sesgos cercanos a cero y azul para sesgos negativos, mientras más intenso sea el tono significa que la magnitud del sesgo es mayor, tanto positiva como negativa. Inicialmente, se observó un sesgo elevado, como se muestra en las figuras 14, 16, 18, 20, 22 y 24, lo cual indica una gran diferencia entre los valores pronosticados y los valores observados. Por otro lado, se realizó un análisis similar para la corrección aplicada por el Filtro de Kalman Adaptativo (figuras 15, 17, 19, 21, 23 y 25), encontrando que el sesgo disminuyó considerablemente, acercándose a cero en las variables de temperatura máxima y temperatura mínima. Para el caso de la precipitación, se muestra una favorable reducción de sesgo, aunque no tan significativa en comparación con la temperatura.

Sesgo pronóstico vs observación - Temperatura máxima (DEF)

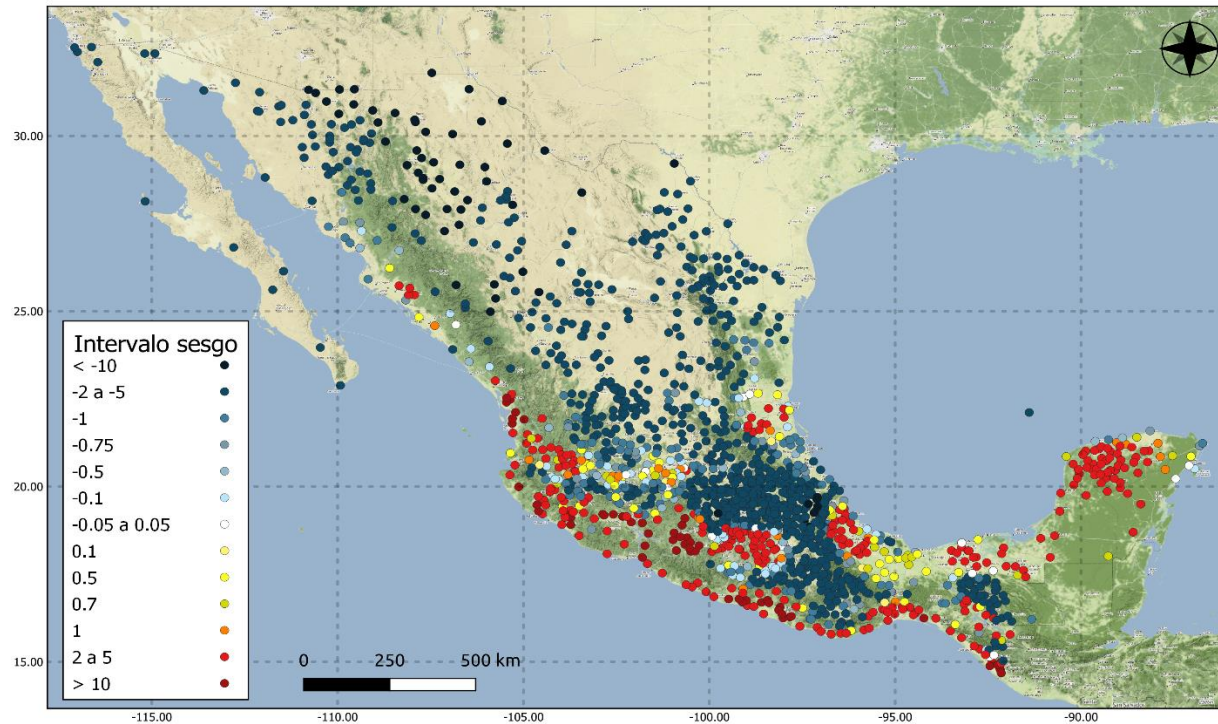


Figura 14: Sesgo de la temperatura máxima entre el pronóstico del SREF y las observaciones de 1570 estaciones climáticas convencionales durante diciembre, enero y febrero de 2020-2021. Representación de sesgos por colores: rojo para sesgos positivos, blanco para sesgos cercanos a cero y azul para sesgos negativos. Elaboración propia mediante el uso de QGIS.

Sesgo corrección vs observación - Temperatura máxima (DEF)

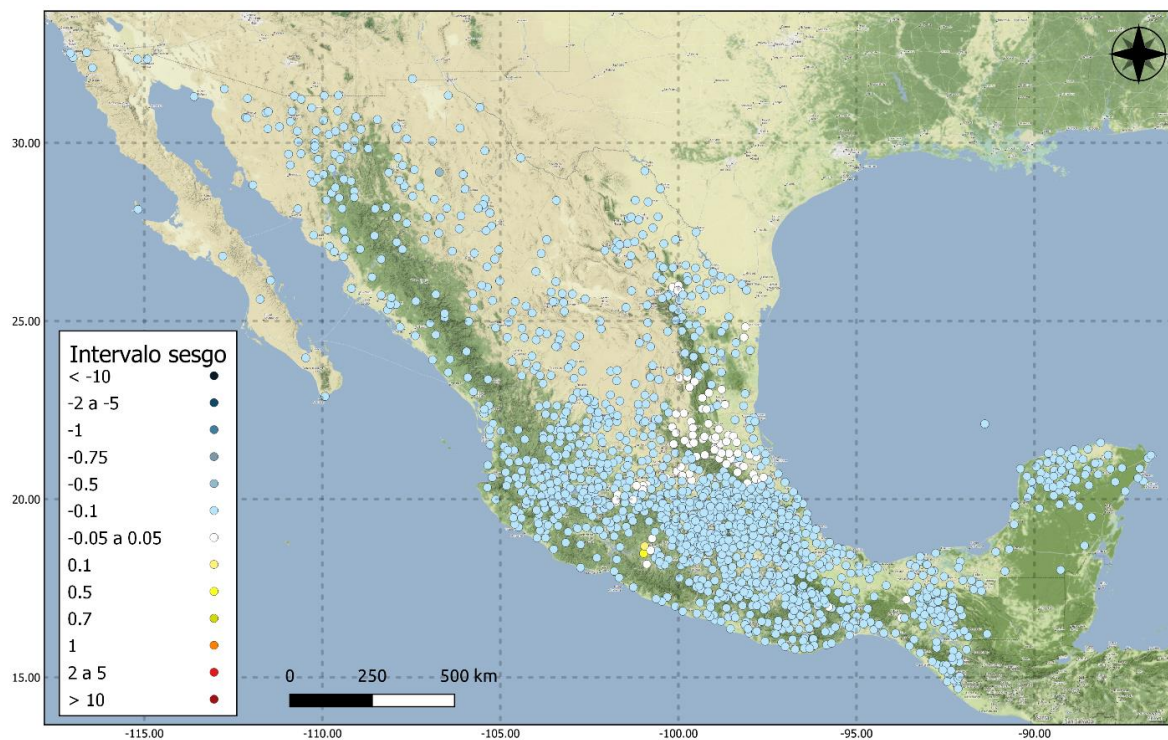


Figura 15: Sesgo de la temperatura máxima entre la corrección mediante Filtro de Kalman Adaptativo y las observaciones de 1570 estaciones climáticas convencionales durante diciembre, enero y febrero de 2020-2021. Representación de sesgos por colores: rojo para sesgos positivos, blanco para sesgos cercanos a cero y azul para sesgos negativos. Elaboración propia mediante el uso de QGIS.

Sesgo pronóstico vs observación - Temperatura máxima (MJJ)

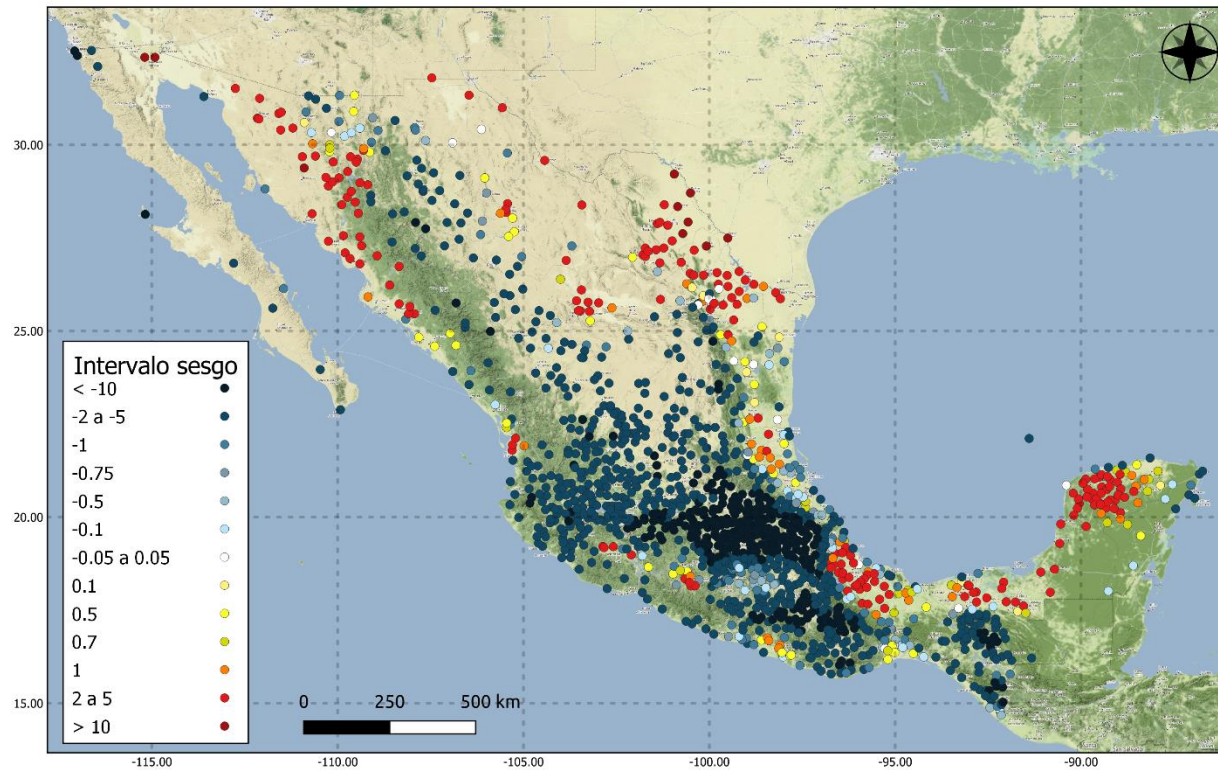


Figura 16: Sesgo de la temperatura máxima entre el pronóstico del SREF y las observaciones de 1570 estaciones climáticas convencionales durante mayo, junio y julio de 2020. Representación de sesgos por colores: rojo para sesgos positivos, blanco para sesgos cercanos a cero y azul para sesgos negativos. Elaboración propia mediante el uso de QGIS

Sesgo corrección vs observación - Temperatura máxima (MJJ)

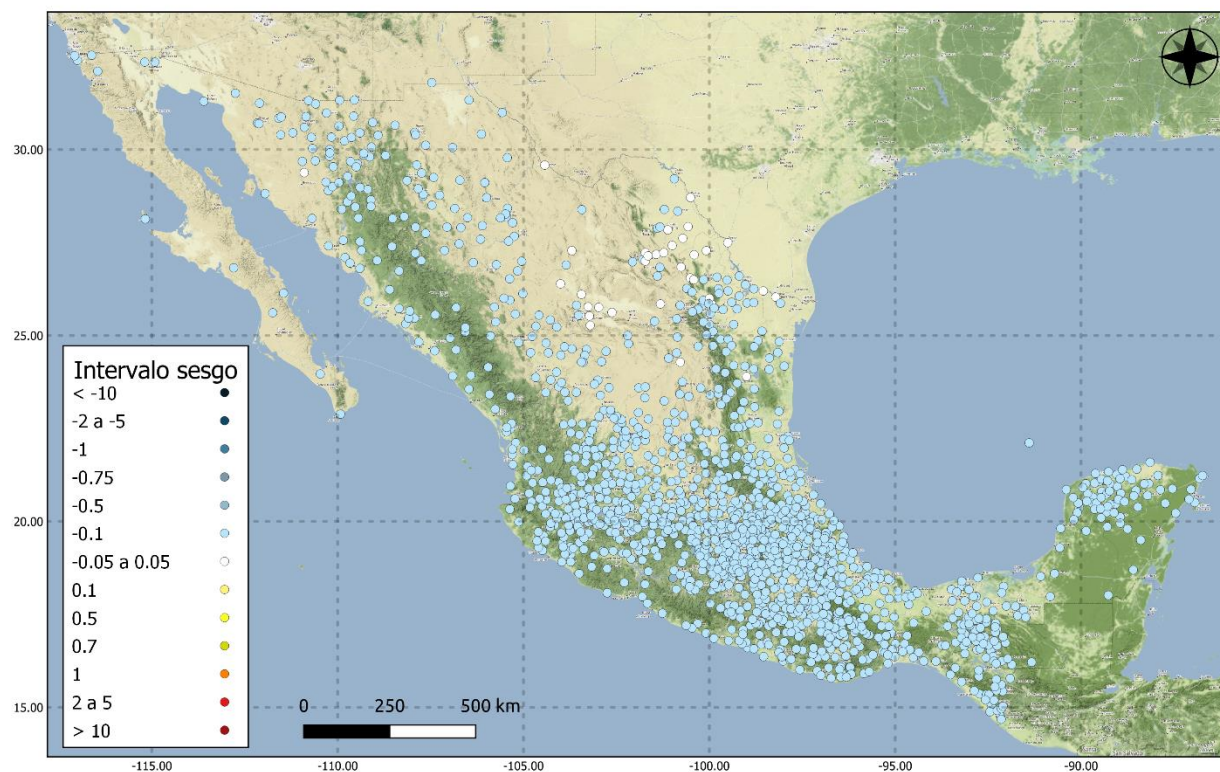


Figura 17: Sesgo de la temperatura máxima entre la corrección mediante Filtro de Kalman Adaptativo y las observaciones de 1570 estaciones climáticas convencionales durante mayo, junio y julio de 2020. Representación de sesgos por colores: rojo para sesgos positivos, blanco para sesgos cercanos a cero y azul para sesgos negativos. Elaboración propia mediante el uso de QGIS.

Sesgo pronóstico vs observación - Temperatura mínima (DEF)

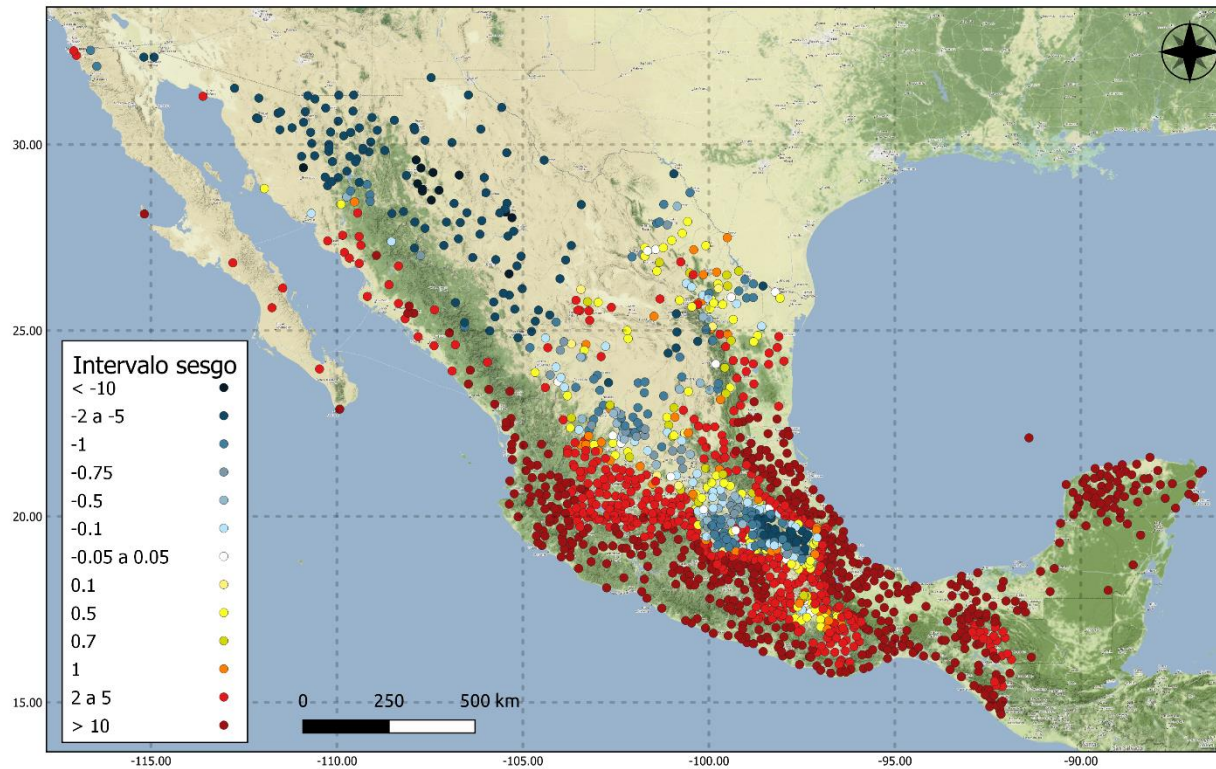


Figura 18: Sesgo de la temperatura mínima entre el pronóstico del SREF y las observaciones de 1570 estaciones climáticas convencionales durante diciembre, enero y febrero de 2020-2021. Representación de sesgos por colores: rojo para sesgos positivos, blanco para sesgos cercanos a cero y azul para sesgos negativos. Elaboración propia mediante el uso de QGIS.

Sesgo corrección vs observación - Temperatura mínima (DEF)

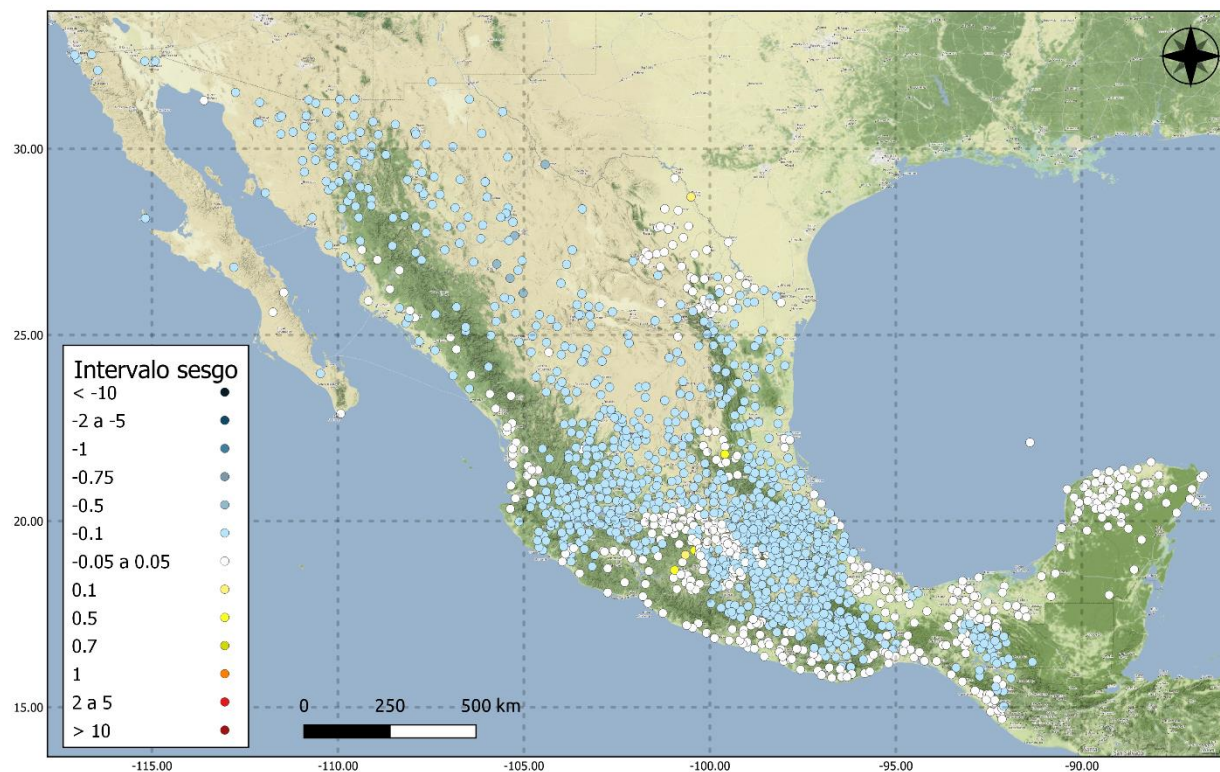


Figura 19: Sesgo de la temperatura mínima entre la corrección mediante Filtro de Kalman Adaptativo y las observaciones de 1570 estaciones climáticas convencionales durante diciembre, enero y febrero de 2020-2021. Representación de sesgos por colores: rojo para sesgos positivos, blanco para sesgos cercanos a cero y azul para sesgos negativos. Elaboración propia mediante el uso de QGIS.

Sesgo pronóstico vs observación - Temperatura mínima (MJJ)

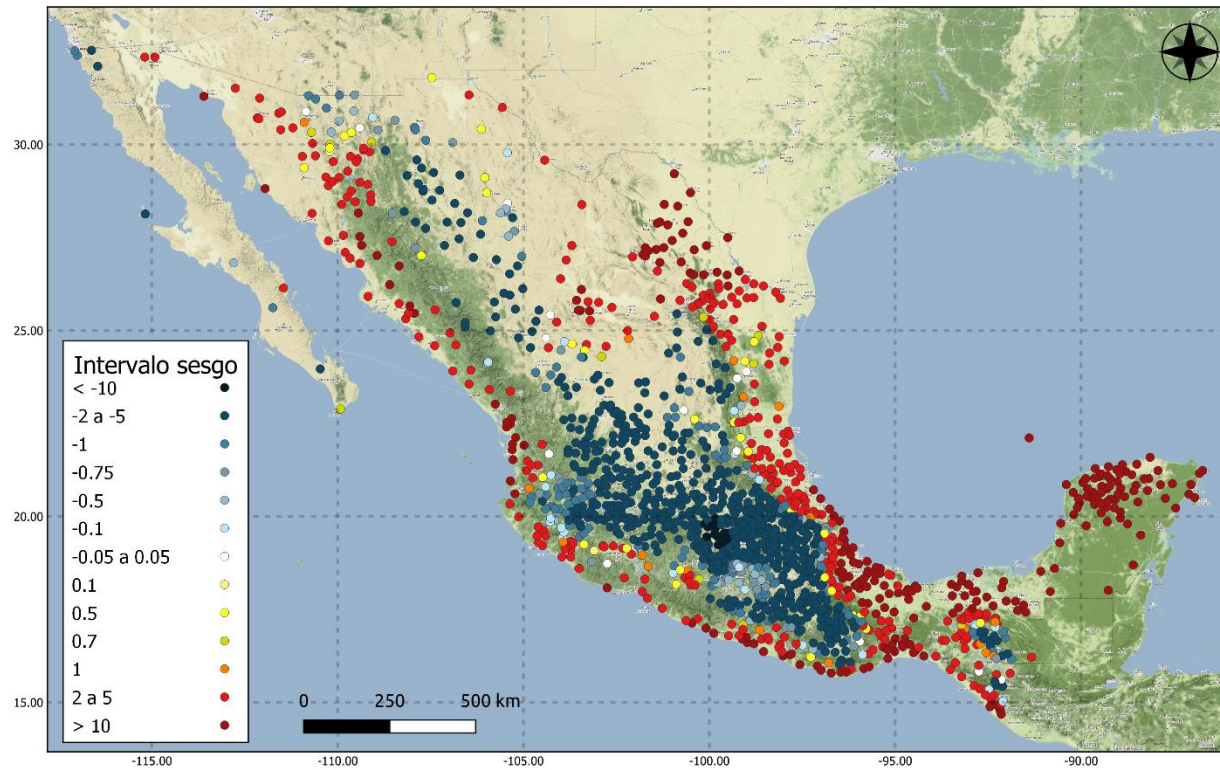


Figura 20: Sesgo de la temperatura mínima entre el pronóstico del SREF y las observaciones de 1570 estaciones climáticas convencionales durante mayo, junio y julio de 2020. Representación de sesgos por colores: rojo para sesgos positivos, blanco para sesgos cercanos a cero y azul para sesgos negativos. Elaboración propia mediante el uso de QGIS.

Sesgo corrección vs observación - Temperatura mínima (MJJ)

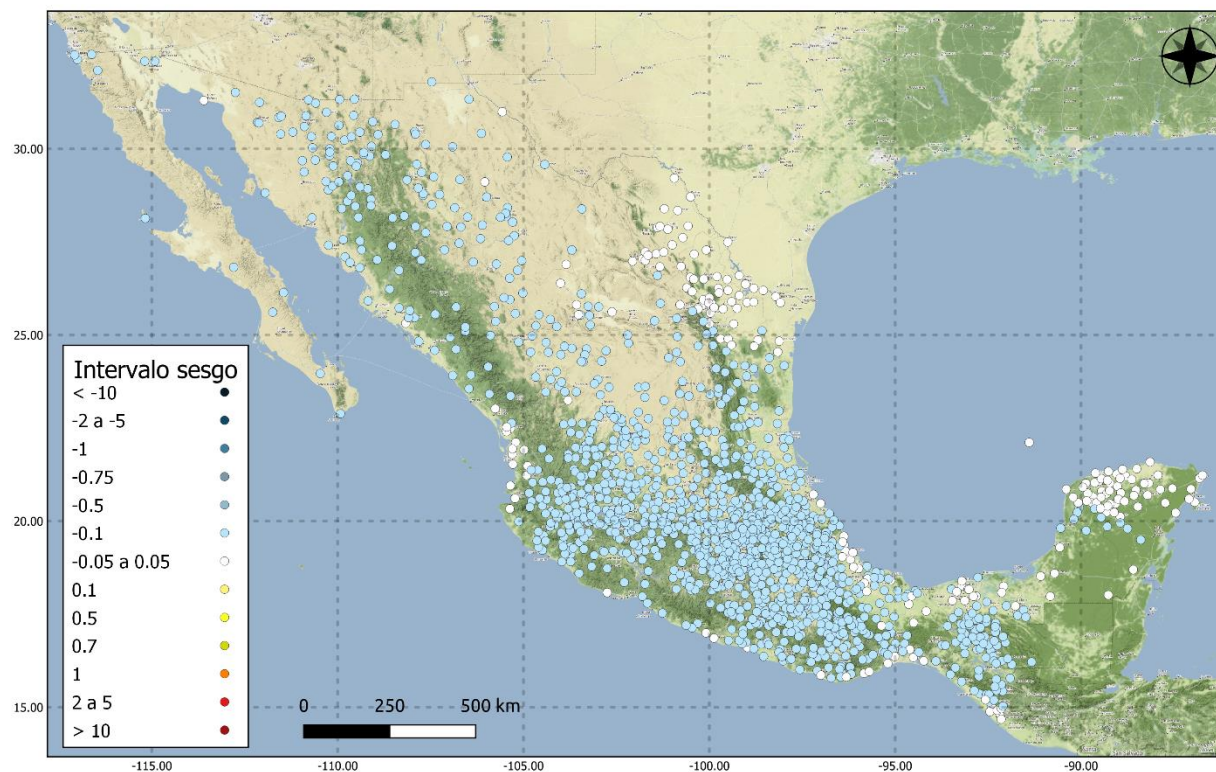


Figura 21: Sesgo de la temperatura mínima entre la corrección mediante Filtro de Kalman Adaptativo y las observaciones de 1570 estaciones climáticas convencionales durante mayo, junio y julio de 2020. Representación de sesgos por colores: rojo para sesgos positivos, blanco para sesgos cercanos a cero y azul para sesgos negativos. Elaboración propia mediante el uso de QGIS.

Sesgo pronóstico vs observación - Precipitación (DEF)

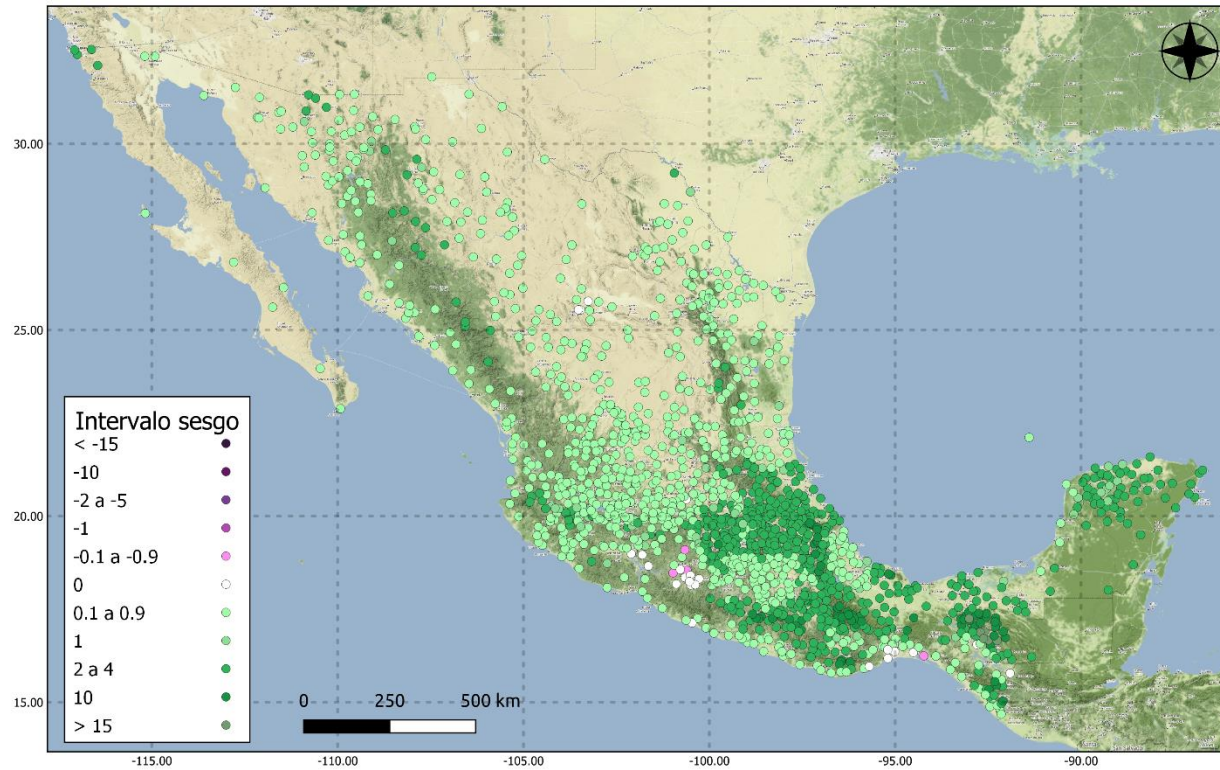


Figura 22: Sesgo de la precipitación entre el pronóstico del SREF y las observaciones de 1570 estaciones climáticas convencionales durante diciembre, enero y febrero de 2020-2021. Representación de sesgos por colores: verdes para sesgos positivos, blanco para sesgos cercanos a cero y morado para sesgos negativos. Elaboración propia mediante el uso de QGIS.

Sesgo corrección vs observación - Precipitación (DEF)

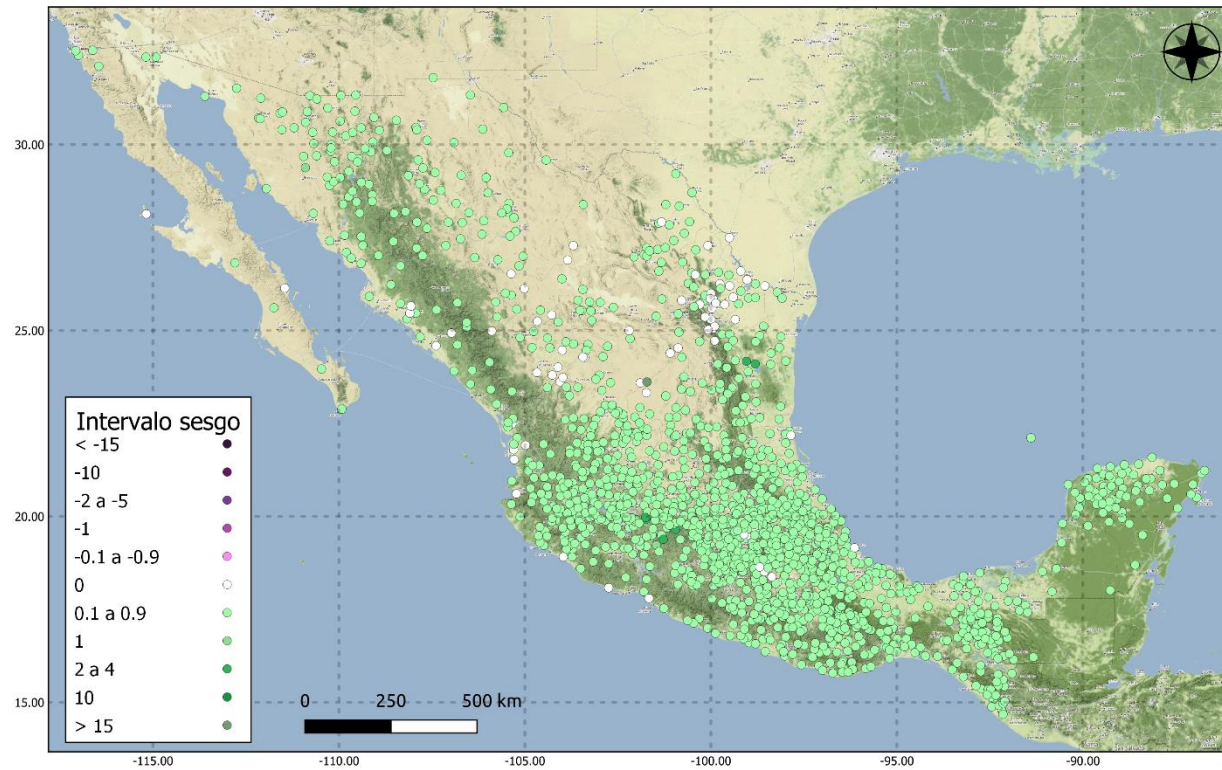


Figura 23: Sesgo de precipitación entre la corrección mediante Filtro de Kalman Adaptativo y las observaciones de 1570 estaciones climáticas convencionales durante diciembre, enero y febrero de 2020-2021. Representación de sesgos por colores: verde para sesgos positivos, blanco para sesgos cercanos a cero y morado para sesgos negativos. Elaboración propia mediante el uso de QGIS.

Sesgo pronóstico vs observación - Precipitación (MJJ)

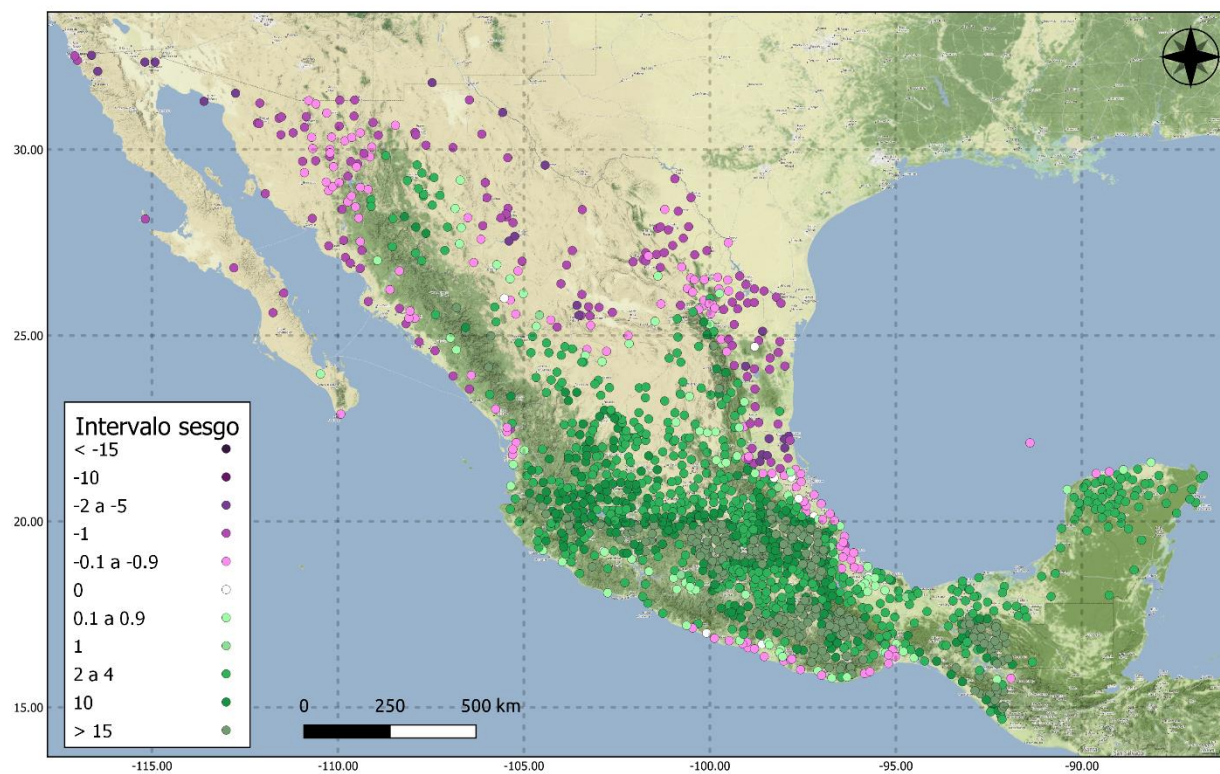


Figura 24: Sesgo de la precipitación entre el pronóstico del SREF y las observaciones de 1570 estaciones climáticas convencionales durante mayo, junio y julio de 2020. Representación de sesgos por colores: verde para sesgos positivos, blanco para sesgos cercanos a cero y morado para sesgos negativos. Elaboración propia mediante el uso de QGIS.

Sesgo corrección vs observación - Precipitación (MJJ)

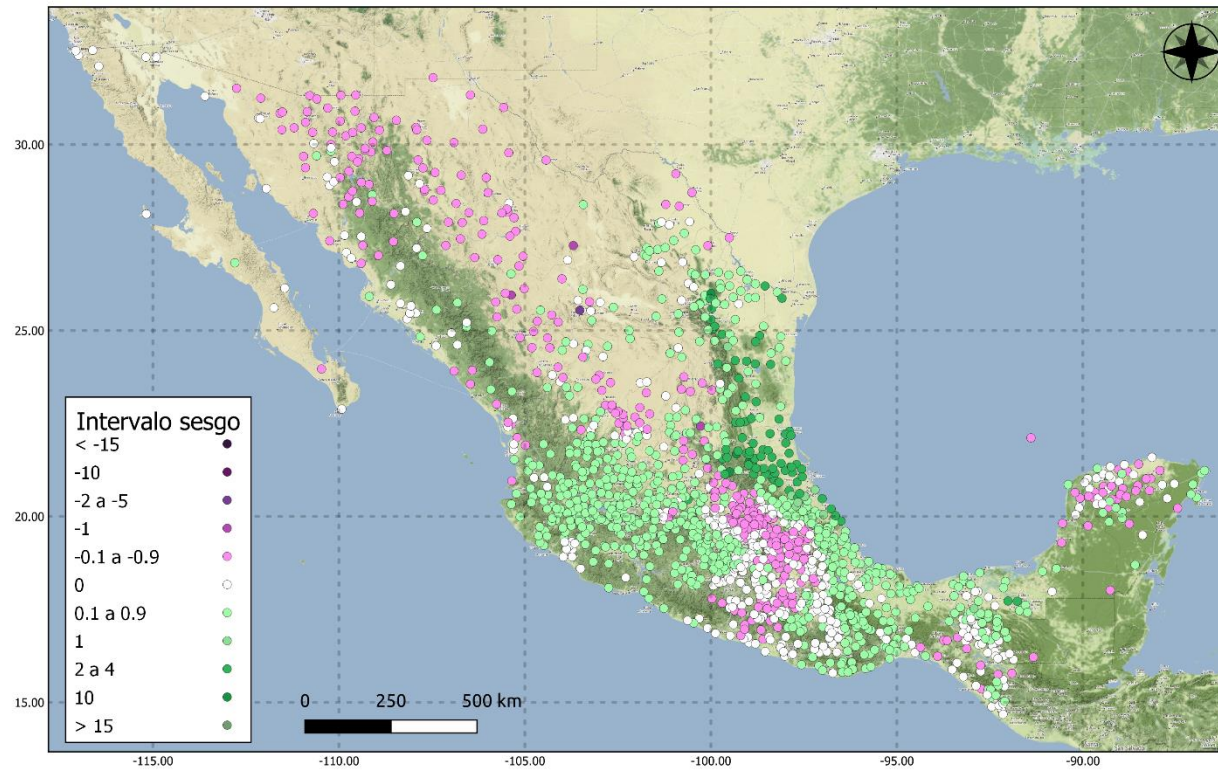


Figura 25: Sesgo de la precipitación entre la corrección mediante Filtro de Kalman Adaptativo y las observaciones de 1570 estaciones climáticas convencionales durante mayo, junio y julio de 2020. Representación de sesgos por colores: verdes para sesgos positivos, blanco para sesgos cercanos a cero y morado para sesgos negativos. Elaboración propia mediante el uso de QGIS.

Variable: TMAX - Periodo: DEF 2020-2021
Municipio: Jiutepec, Estado: Mor.

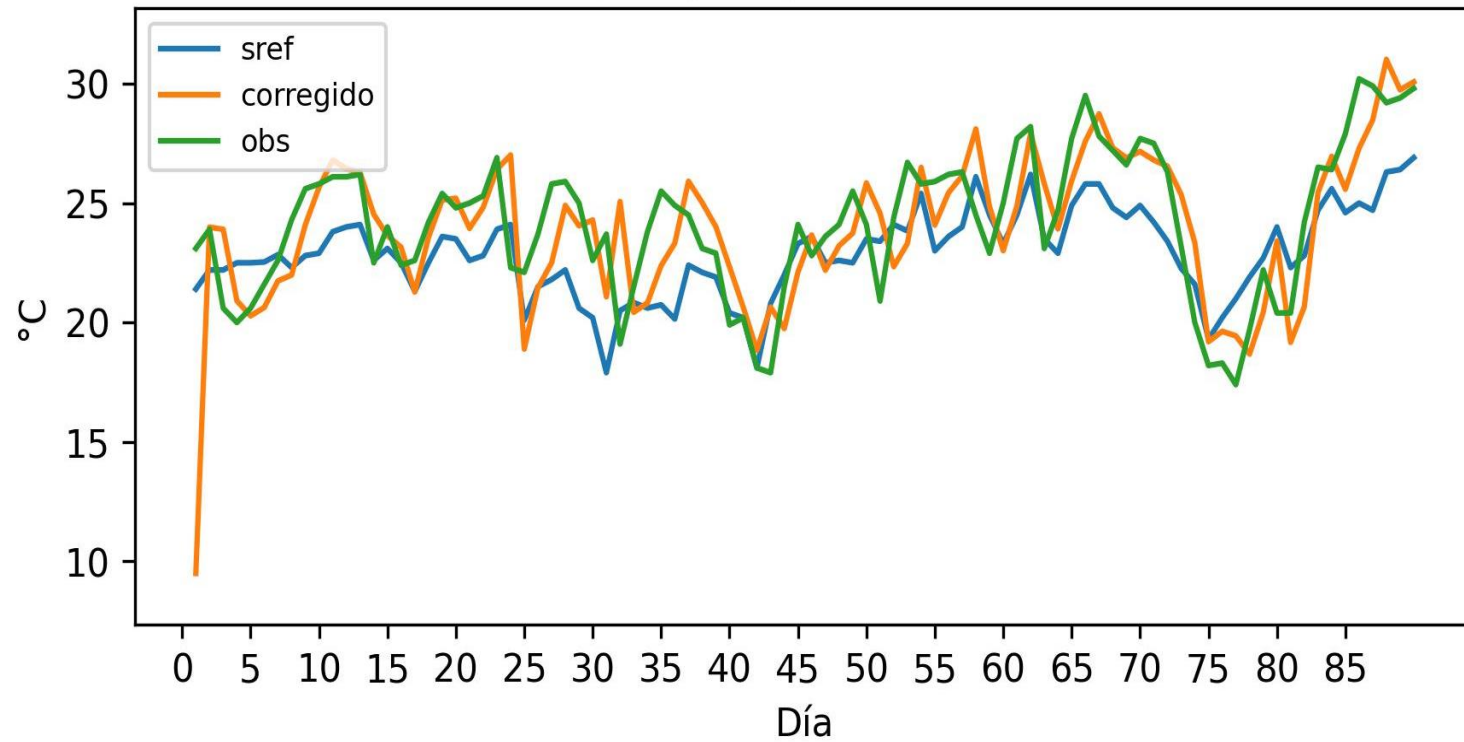


Figura 26: Comparación entre datos de temperatura máxima extraídos de la media del ensamble SREF (azul) para las coordenadas de la estación de Jiutepec, Mor., temperaturas máximas corregidas (naranja) con Filtro de Kalman Adaptado y observaciones (verde), periodo DEF.

Variable: TMAX - Periodo: MJJ 2020
Municipio: Jiutepec, Estado: Mor.

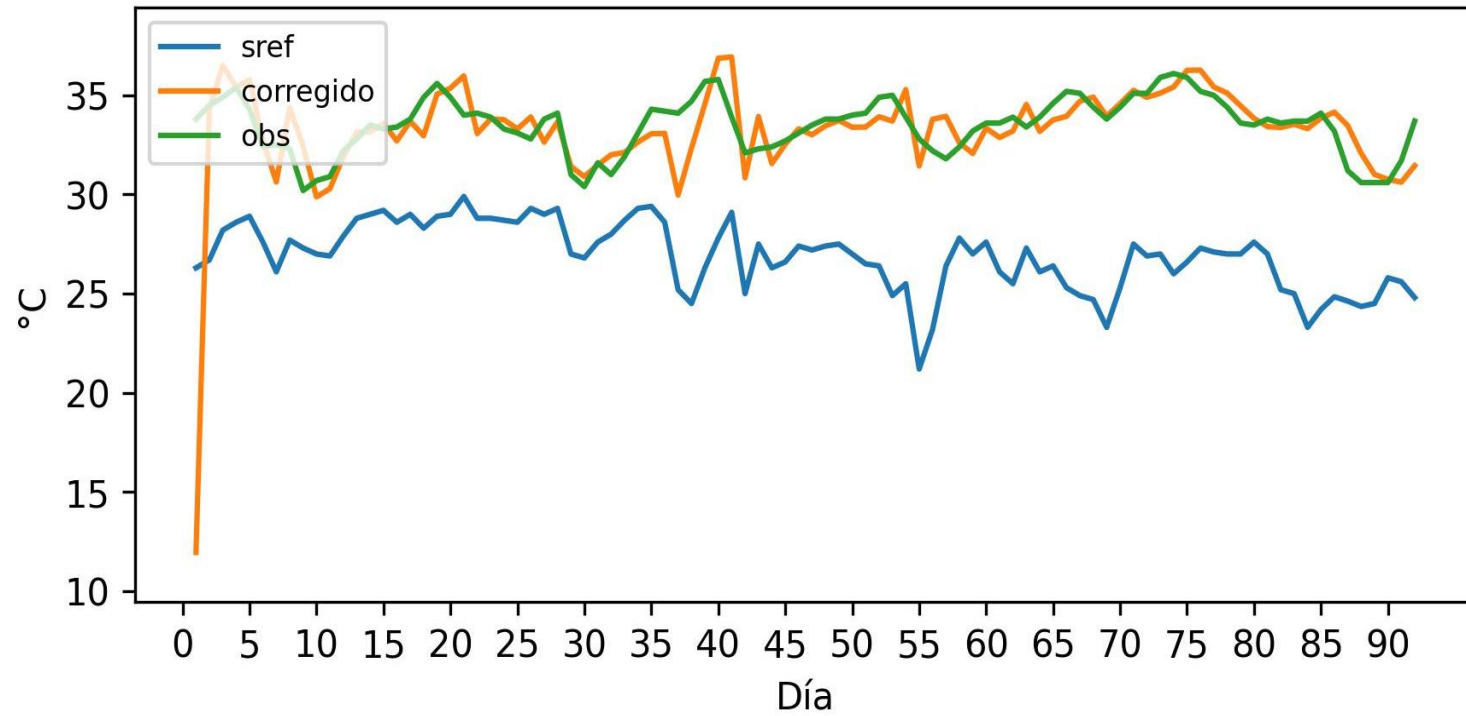


Figura 27: Comparación entre datos de temperatura máxima extraídos de la media del ensamble SREF (azul) para las coordenadas de la estación de Jiutepec, Mor., temperaturas máximas corregidas (naranja) con Filtro de Kalman Adaptado y observaciones (verde), periodo MJJ.

Variable: TMIN - Periodo: DEF 2020-2021
Municipio: Jiutepec, Estado: Mor.

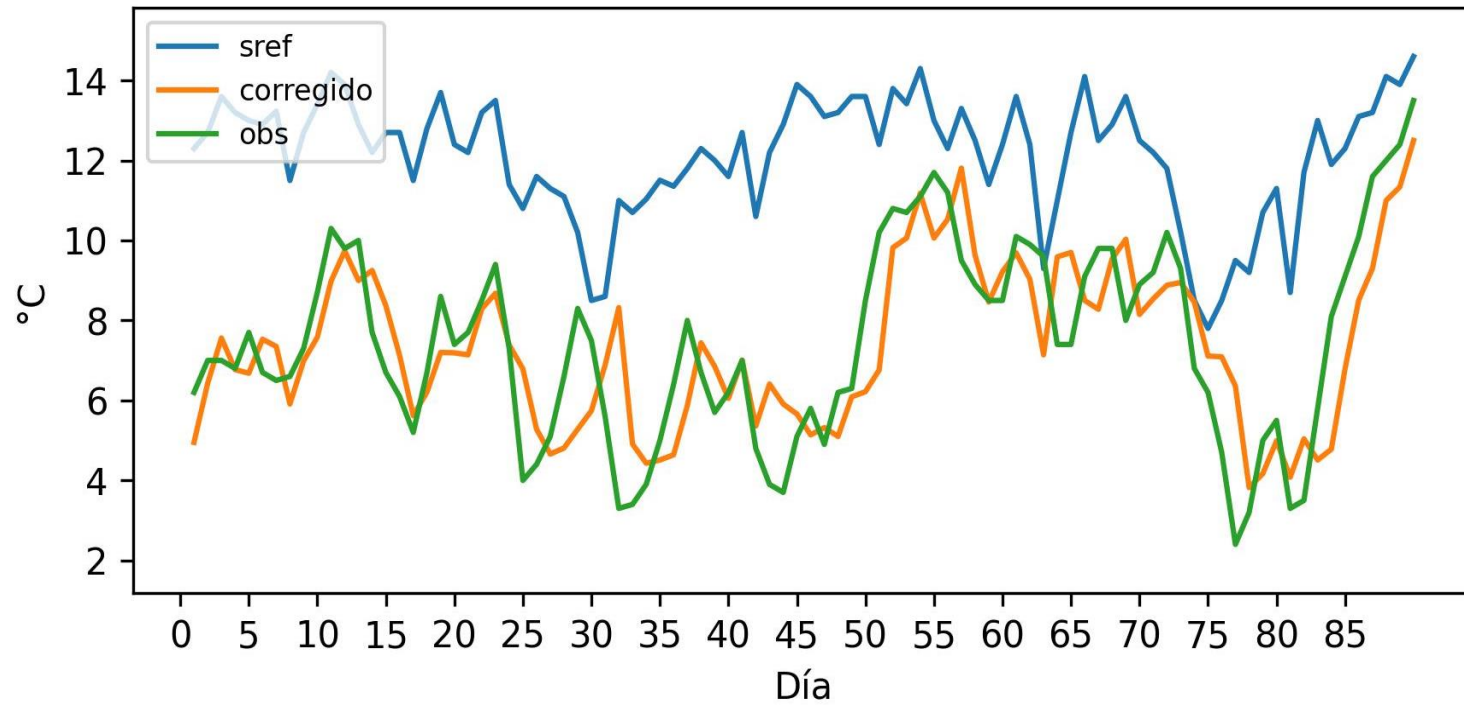


Figura 28: Comparación entre datos de temperatura mínima extraídos de la media del ensamble SREF (azul) para las coordenadas de la estación de Jiutepec, Mor., temperaturas máximas corregidas (naranja) con Filtro de Kalman Adaptado y observaciones (verde), periodo DEF.

Variable: TMIN - Periodo: MJJ 2020
Municipio: Jiutepec, Estado: Mor.

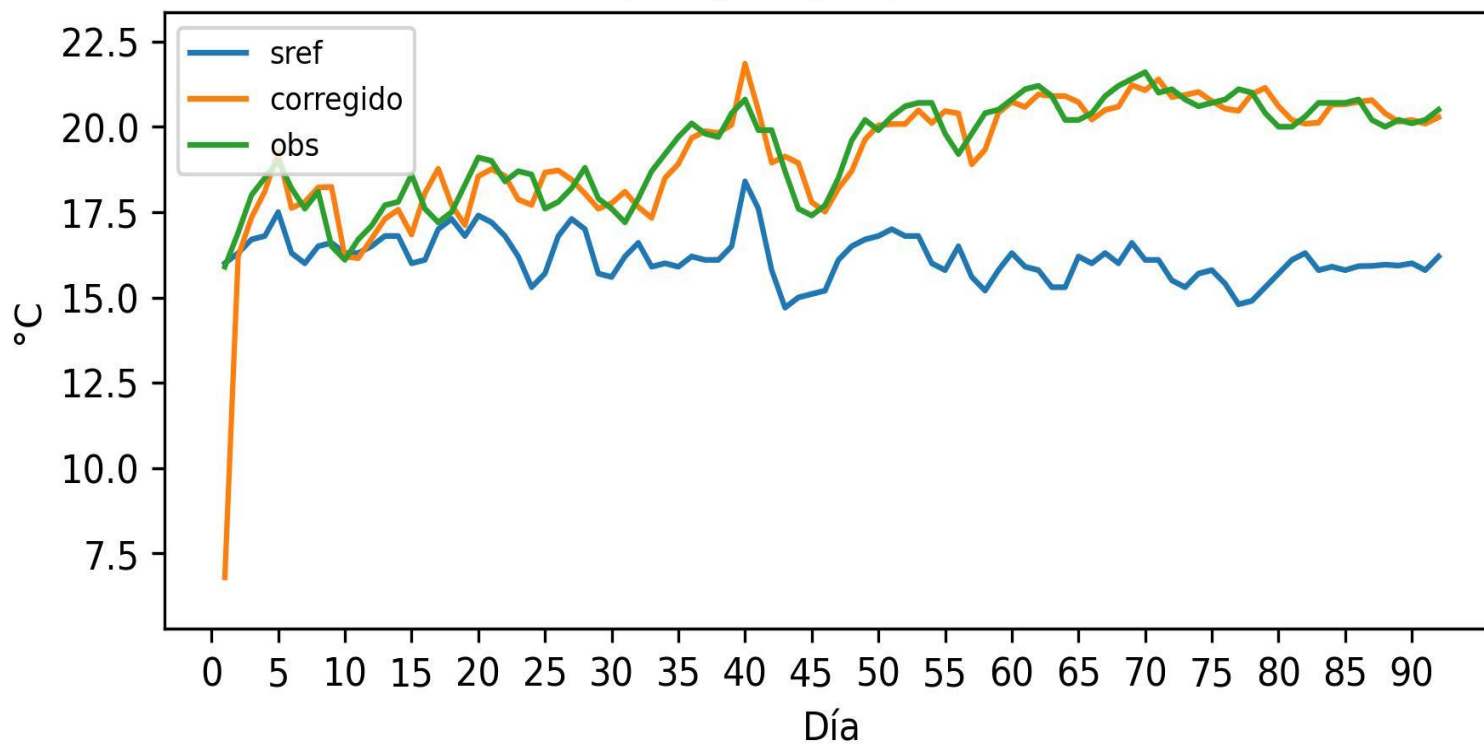


Figura 29: Comparación entre datos de temperatura mínima extraídos de la media del ensemble SREF (azul) para las coordenadas de la estación de Jiutepec, Mor., temperaturas máximas corregidas (naranja) con Filtro de Kalman Adaptado y observaciones (verde), periodo MJJ.

Variable: PRECIPITACIÓN - Periodo: DEF 2020-2021
Municipio: Jiutepec, Estado: Mor.

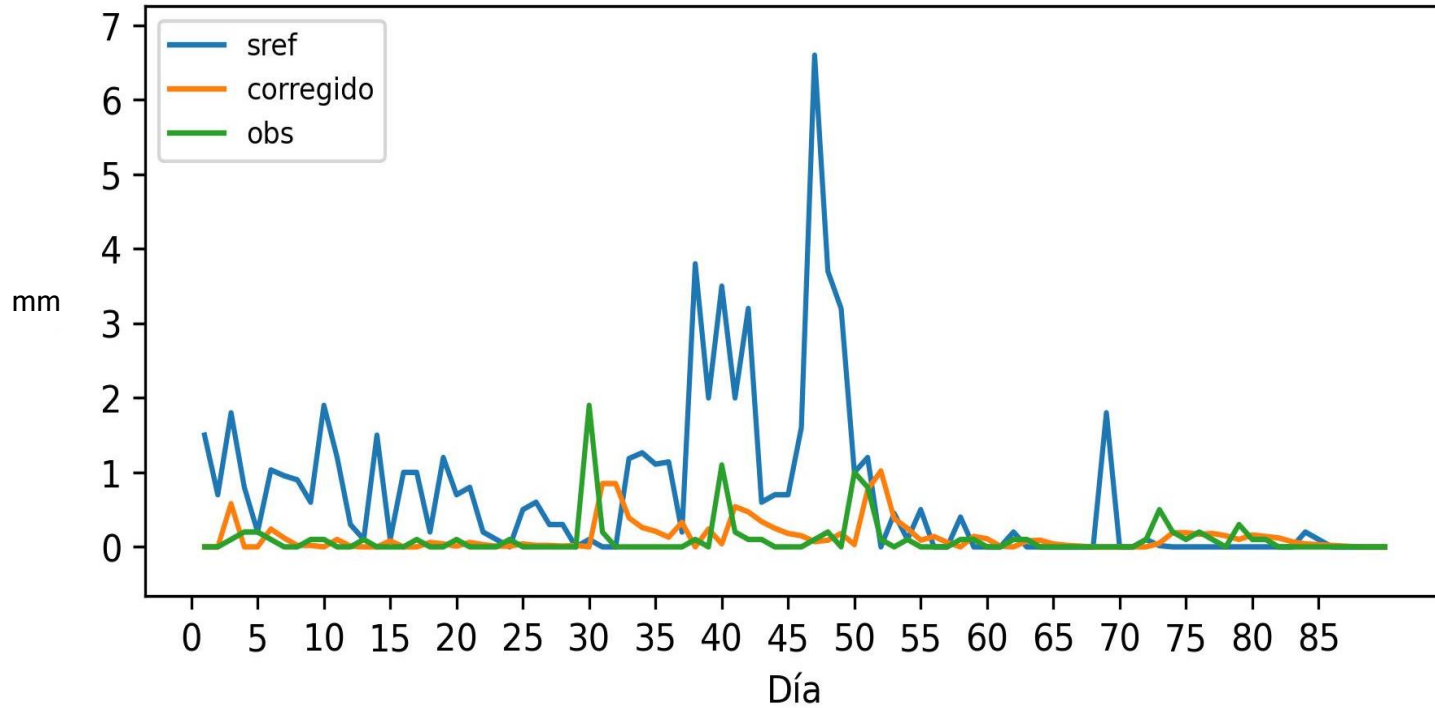


Figura 30: Comparación entre datos de precipitación extraídos de la media del ensamble SREF (azul) para las coordenadas de la estación de Jiutepec, Mor., precipitaciones corregidas (naranja) con Filtro de Kalman Adaptado y observaciones (verde), periodo DEF.

Variable: PRECIPITACIÓN - Periodo: MJJ 2020
Municipio: Jiutepec, Estado: Mor.

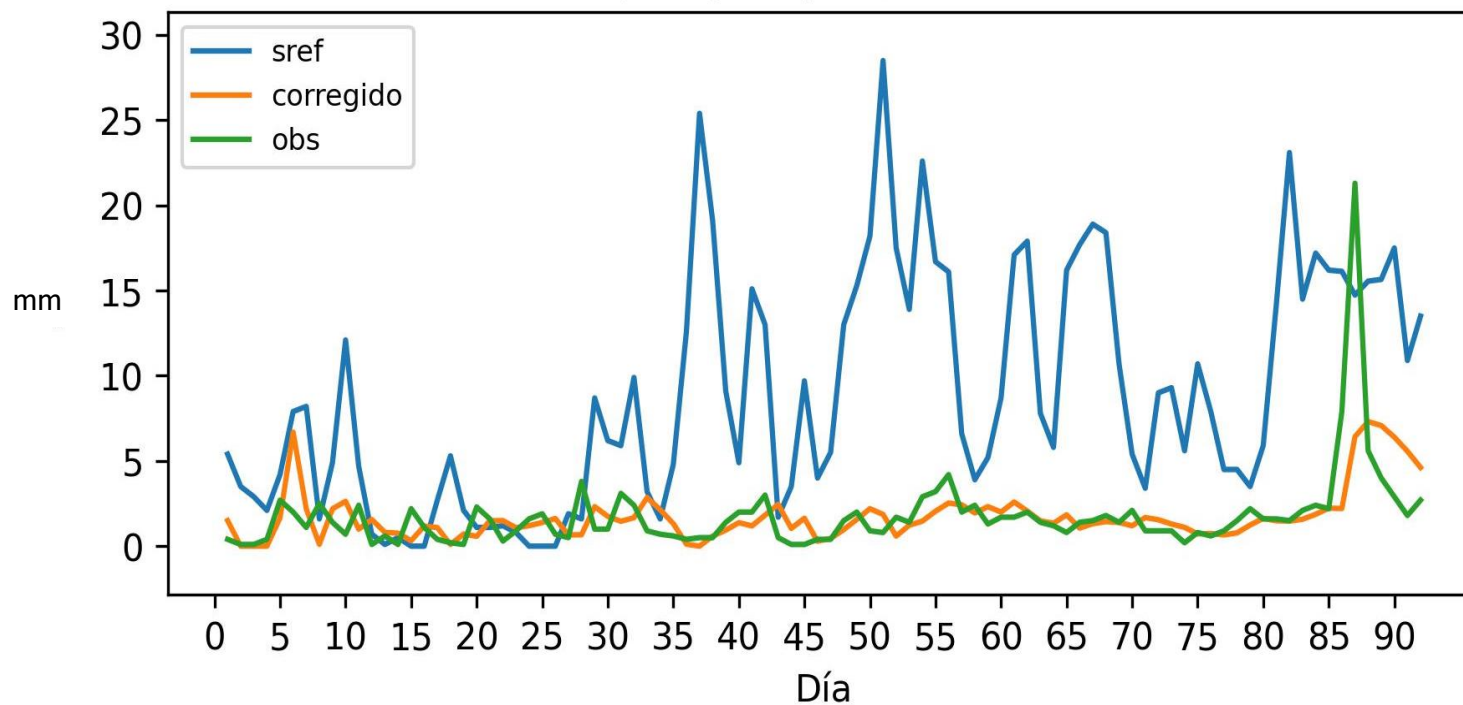


Figura 31: Comparación entre datos de precipitación extraídos de la media del ensamble SREF (azul) para las coordenadas de la estación de Jiutepec, Mor., precipitaciones corregidas (naranja) con Filtro de Kalman Adaptado y observaciones (verde), periodo MJJ.

A manera de ejemplo, se muestran series temporales de los trimestres analizados para cada una de las variables (Figuras 26, 27, 28, 29, 30, 31), estas figuras ilustran la variable pronosticada por SREF en color azul, la variable corregida con el Filtro de Kalman Adaptativo en color naranja y las observaciones registradas en color verde para la estación de Jiutepec, Morelos. Se logra apreciar, que, a partir del segundo día de análisis, una vez que los coeficientes de corrección son actualizados, el filtro logra corregir los pronósticos del sistema SREF, acercándose considerablemente a las observaciones registradas para cada día.

Identificación de reducción de sesgo al aplicar Filtro de Kalman Adaptativo

En esta sección, se examinará la efectividad de la reducción de sesgo al utilizar el Filtro de Kalman Adaptativo en las simulaciones numéricas del sistema SREF. Siguiendo un enfoque similar al utilizado en la sección 2.4 para identificar el sesgo, se aplicaron las mismas métricas estadísticas y se tomaron los mismos casos de estudio. El objetivo fue comparar los resultados entre el pronóstico bruto proporcionado por el sistema SREF y el pronóstico corregido por el Filtro de Kalman Adaptativo, a fin de evaluar el rendimiento de la metodología.

Los resultados de estas métricas estadísticas se presentan en tablas que muestran los intervalos de sesgo, agrupados en rangos predefinidos. Los resultados obtenidos de esta sección para el RMSE, MAE y COE se visualizan en el Anexo 3, que se presentan más adelante. Cada intervalo muestra el número de datos y el porcentaje correspondiente del total de datos que caen dentro de ese intervalo de sesgo. Estas tablas resumen la distribución del sesgo en las diferentes estaciones, lo que permite una visualización y análisis del sesgo en función de los distintos rangos de valores.

A continuación, se presentan las tablas y figuras correspondientes a los intervalos de sesgo de temperatura máxima, temperatura mínima y precipitación para los trimestres DEF 2020-2021 y MJJ 2020. Los productos muestran el sesgo entre el pronóstico corregido a través de Filtro de Kalman Adaptativo y las observaciones de estaciones climáticas convencionales:

Tabla 11: Intervalos de sesgo para la temperatura máxima en el trimestre de diciembre, enero y febrero 2020-2021, pronóstico vs observación.

INTERVALO	NUM. DATOS	%
[-2.0, 0.0]	1562	99.4904459
[0.0, 2.0]	8	0.50955414

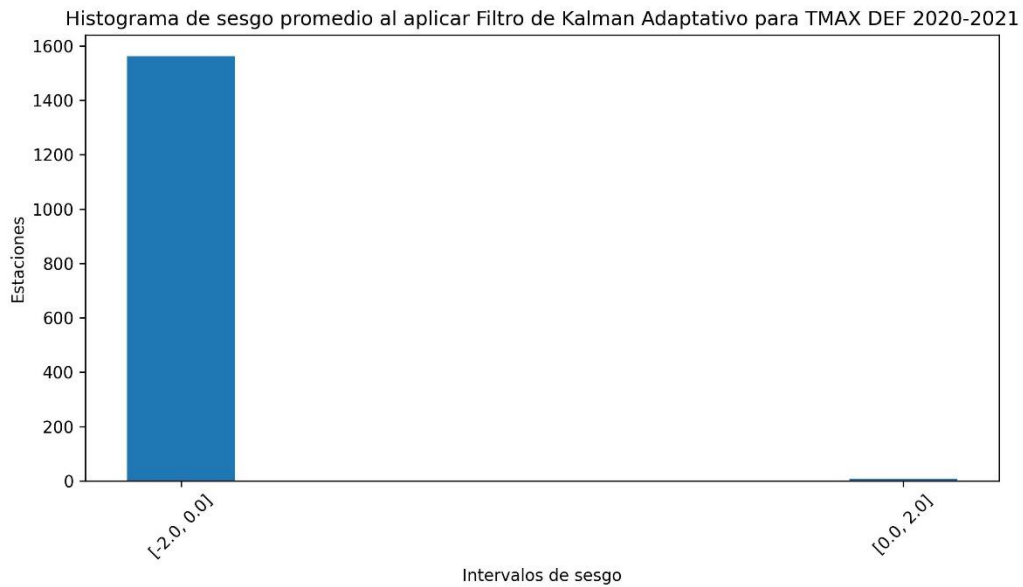


Figura 32: Histograma de sesgo promedio para la temperatura máxima en el periodo DEF 2020-2021, corrección vs observación.

Tabla 12: Intervalos de sesgo para la temperatura máxima en el trimestre de mayo, junio y julio 2020, pronóstico vs observación.

INTERVALO	NUM. DATOS	%
[-2.0, 0.0]	1569	99.9363057
[0.0, 2.0]	1	0.06369427

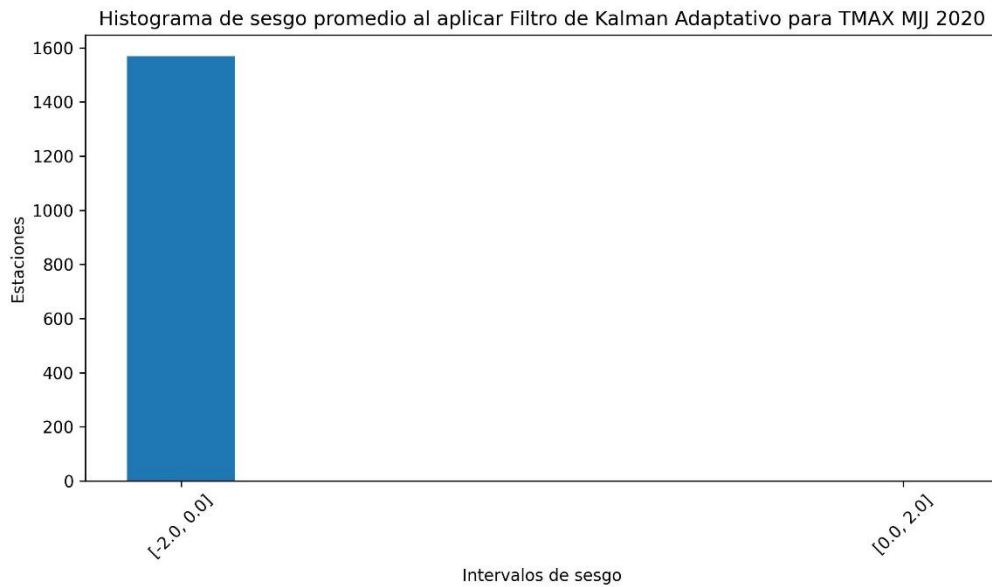


Figura 33: Histograma de sesgo promedio para la temperatura máxima en el periodo MJJ 2020, corrección vs observación.

Tabla 13: Intervalos de sesgo para la temperatura mínima en el trimestre de diciembre, enero y febrero 2020-2021, pronóstico vs observación.

INTERVALO	NUM. DATOS	%
[-2.0, 0.0]	1548	98.5987261
[0.0, 2.0]	22	1.40127389

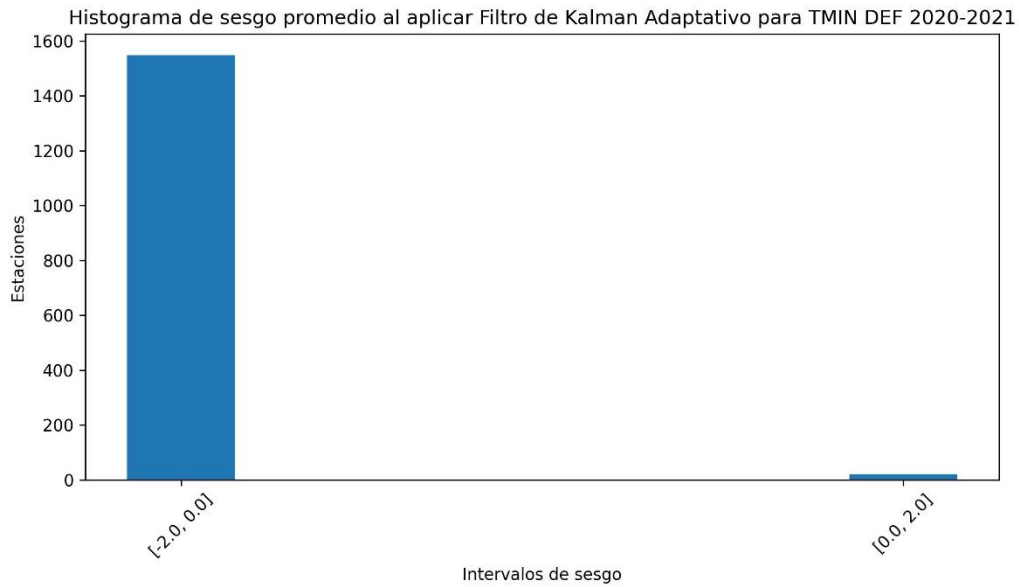


Figura 34: Histograma de sesgo promedio para la temperatura mínima en el periodo DEF 2020-2021, corrección vs observación.

Tabla 14: Intervalos de sesgo para la temperatura mínima en el trimestre de mayo, junio y julio 2020, corrección vs observación.

INTERVALO	NUM. DATOS	%
[-2.0, 0.0]	1570	100

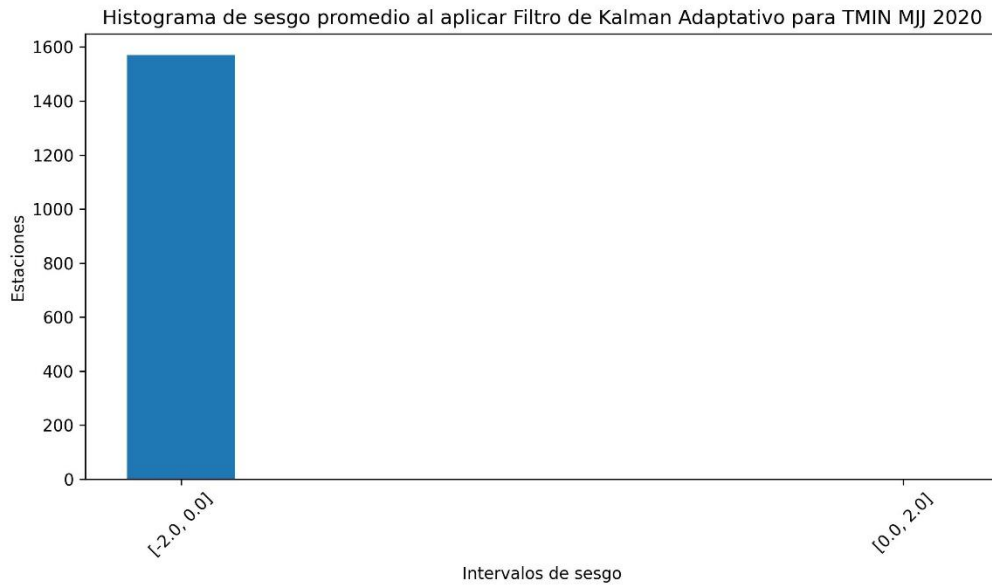


Figura 35: Histograma de sesgo promedio para la temperatura mínima en el periodo MJJ 2020, corrección vs observación.

Tabla 15: Intervalos de sesgo para la precipitación en el trimestre de diciembre, enero y febrero 2020-2021, corrección vs observación.

INTERVALO	NUM. DATOS	%
[-2.0, 0.0]	71	4.52229299
[0.0, 2.0]	1498	95.4140127
...
[32.0, 34.0]	1	0.06369427

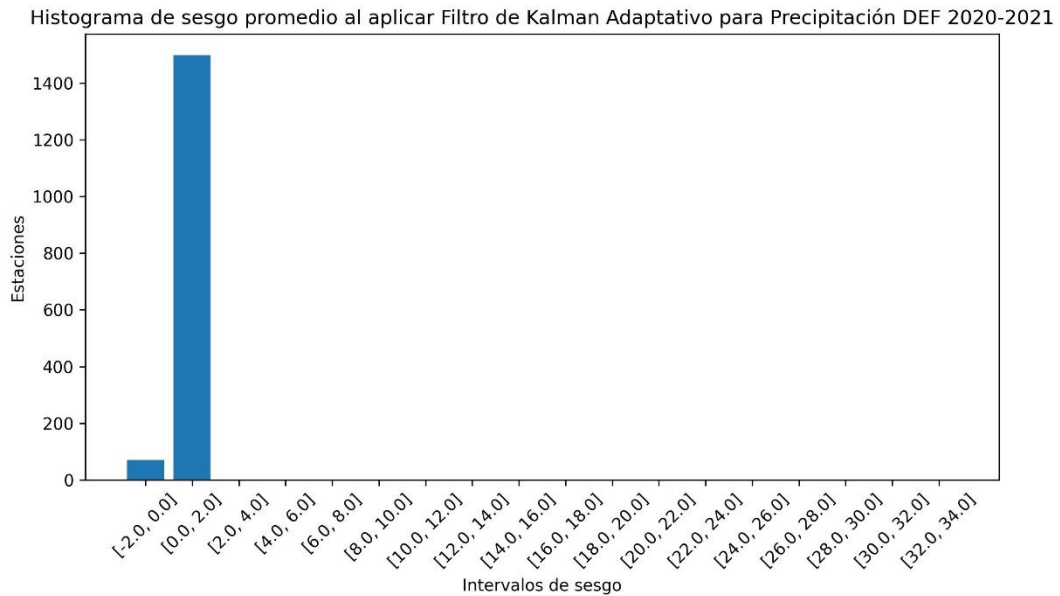


Figura 36: Histograma de sesgo promedio para la Precipitación en el periodo DEF 2020-2021, corrección vs observación.

Tabla 16: Intervalos de sesgo para la precipitación en el trimestre de mayo, junio y julio 2020, corrección vs observación.

INTERVALO	NUM. DATOS	%
[-4.0, -2.0]	1	0.06369427
[-2.0, 0.0]	756	48.1528662
[0.0, 2.0]	785	50
[2.0, 4.0]	22	1.40127389
[4.0, 6.0]	3	0.1910828
[6.0, 8.0]	1	0.06369427
[8.0, 10.0]	2	0.12738854

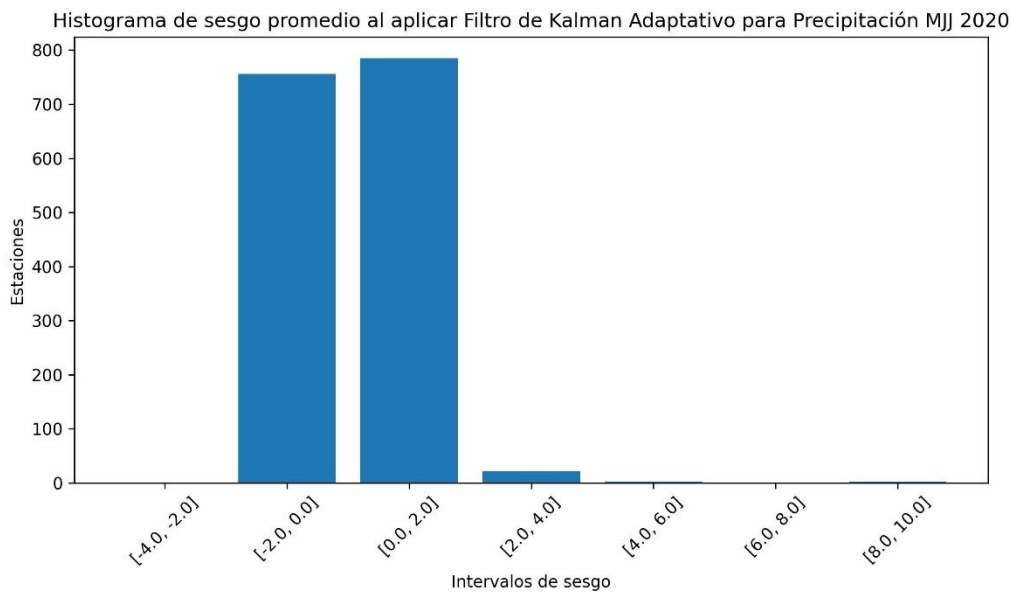


Figura 37: Histograma de sesgo promedio para la Precipitación en el periodo MJJ 2020, corrección vs observación.

Al analizar las tablas 11, 12, 13 y 14 correspondientes a las variables de temperatura máxima y mínima para los trimestres DEF y MJJ, y las Figuras 32, 33, 34 y 35, se observa que el sesgo identificado se encuentra dentro del rango de $[-2$ a $2]$, con la excepción de la temperatura mínima en MJJ, donde el 100 % de las estaciones presentan un sesgo negativo. Con el fin de realizar una comparación con las tablas presentadas en la sección “Aplicación de métodos estadísticos para identificación de sesgo”, se optó por establecer intervalos de dos unidades. Sin embargo, al examinar detalladamente la tabla 17, se evidencia de manera significativa que el sesgo identificado es extremadamente próximo a cero. Esto sugiere que el Filtro de Kalman Adaptativo ha logrado reducir eficientemente el sesgo presente en los resultados de las simulaciones numéricas del sistema SREF. La proximidad al valor cero indica que el pronóstico corregido se acerca de manera adecuada a los valores observados o de referencia, lo cual es un indicio positivo de la precisión y confiabilidad de la metodología empleada. Este resultado prometedor subraya el potencial del Filtro de Kalman Adaptativo como una herramienta eficaz para mitigar el sesgo en pronósticos meteorológicos numéricos y mejorar la calidad de los resultados.

Se aprecia una mejoría en los resultados de la variable de precipitación, como se puede observar en la tabla 15 – Figura 36, y tabla 16 – Figura 37, aunque esta mejoría no es tan notable como en el caso de la temperatura. Esto podría atribuirse a las consideraciones que se plantearon en la sección anterior al aplicar el filtro a esta variable. Aunque las mejoras no fueron tan significativas como en el caso de la temperatura, demuestran el beneficio de implementar el Filtro de Kalman Adaptativo en la corrección de sesgos en pronósticos de precipitación.

Tabla 17: Valor de máximo sesgo negativo y positivo encontrado para cada caso de estudio, corrección vs observación.

Variable - periodo	Máximo sesgo positivo encontrado	Máximo sesgo negativo encontrado
Temperatura máxima - DEF	0.24	-0.59
Temperatura máxima - MJJ	0.04	-0.55
Temperatura mínima - DEF	0.15	-0.69
Temperatura mínima - MJJ	-0.02	-0.3
Precipitación - DEF	33.73	-0.09
Precipitación - MJJ	8.56	-3.15

En la tabla 18, presenta una descripción general de lo identificado en los resultados anteriores:

Tabla 18: Resumen de sesgos e intervalos para las variables de interés, corrección vs observación

Variable	Intervalo promedio	Sesgo promedio	Sesgo absoluto promedio
Temperatura	[-2 a 2]	-0.18 °C	0.18 °C
Temperatura máxima	[-2 a 2]	-0.21 °C	0.21 °C
Temperatura mínima	[-2 a 0]	-0.14 °C	0.14 °C
Temperatura (DEF)	[-2 a 2]	-0.16 °C	0.16 °C
Temperatura (MJJ)	[-2 a 2]	-0.19 °C	0.19 °C
Precipitación	[-3 a 22]	0.11 mm	0.17 mm
Precipitación (DEF)	-	0.09 mm	0.09 mm
Precipitación (MJJ)	-	0.13 mm	0.25 mm

Como complemento, se realizó el mismo enfoque utilizado en Velásquez et al. (2020), y se encontró que el sesgo promedio para la variable de precipitación en los meses de DEF pasó de 0.09 mm a -0.14 mm, mientras que el sesgo absoluto promedio pasó de 0.09 mm a 0.39 mm. Para los meses de MJJ, el sesgo promedio pasó de 0.13 mm a 0.07 mm, y para el sesgo absoluto promedio, encontramos un aumento de 0.25 mm a 0.31 mm.

Tabla 19: Distribución del sesgo positivo y negativo para cada una de las variables y periodos analizados, identificación del comportamiento de la corrección por parte del Filtro de Kalman Adaptativo.

Variable - periodo	Número de datos con sesgo < 0	Número de datos con sesgo > 0	Porcentaje de datos con sesgo < 0	Porcentaje de datos con sesgo > 0	Comportamiento de la corrección
Temperatura máxima - DEF	1562	8	99.49	0.5	Subestima
Temperatura máxima - MJJ	1569	1	99.93	0.06	Subestima
Temperatura mínima - DEF	1548	22	98.59	1.4	Subestima
Temperatura mínima - MJJ	1570	0	100	0	Subestima
Precipitación - DEF	71	1499	4.52	95.47	Sobrestima
Precipitación - MJJ	757	813	48.21	51.78	Sobrestima

La tabla 19 presenta la distribución de sesgos y tendencias observadas mediante la corrección con el Filtro de Kalman Adaptativo para las variables y periodos considerados hasta ahora. Siguiendo una estructura similar a la tabla 10, la tabla 19 muestra la cantidad de datos con sesgos inferiores a cero y superiores a cero para cada variable y periodo, así como el porcentaje correspondiente de cada grupo en relación al total de datos. Además, se identifica la tendencia del modelo (subestimación o sobreestimación) para cada combinación de variable y periodo. Esta tabla proporciona una visión precisa del comportamiento de la tendencia de la corrección en términos de sobreestimar o subestimar las diferentes variables en diferentes épocas del año. Estos hallazgos podrían tenerse en cuenta al generar pronósticos meteorológicos para la población, permitiendo una mayor precisión y ajuste en las predicciones.

De acuerdo con la información presentada en la tabla 19, se observa que la corrección realizada subestima las variables de temperatura y sobrestima la precipitación en todos los casos analizados. Esto indica que los valores observados en el tiempo t son más altos que el pronóstico corregido en el tiempo $t-1$. Para ejemplificar esto, consideremos un caso hipotético: supongamos que se emite un pronóstico corregido por SREF para el día de mañana, y se obtiene una temperatura máxima de 25.18 °C en una estación del país. Teniendo en cuenta que el sesgo promedio para la variable de temperatura es de -0.18 °C, es probable que la temperatura máxima real que se presente el día siguiente sea de aproximadamente 25 °C.

Es importante resaltar que este ejemplo hipotético se presenta a modo de explicación. Además, cabe destacar que el pronóstico corregido demuestra ser muy preciso. Al mencionar que la corrección subestima, nos referimos a que los valores corregidos son más bajos que los observados, aunque la magnitud sea muy pequeña.

El análisis realizado en esta sección ha evaluado la efectividad de la reducción de sesgo al utilizar el Filtro de Kalman Adaptativo en las simulaciones numéricas del sistema SREF. Los resultados obtenidos demuestran que esta metodología ha logrado reducir eficientemente el sesgo presente en los pronósticos del sistema SREF. Se ha observado una mejora significativa en la precisión de las variables de temperatura, donde el sesgo identificado se encuentra extremadamente próximo a cero, indicando una mayor aproximación a los valores observados. Aunque las mejoras en la variable de precipitación no son tan notables, se evidencia un beneficio al aplicar el Filtro de Kalman Adaptativo para esta variable, lo que sugiere seguir esta línea de investigación y realizar las mejoras correspondientes.

En el siguiente capítulo, se presentarán las conclusiones obtenidas a partir de este estudio, destacando los principales resultados, implicaciones y recomendaciones para futuras investigaciones en el campo de la corrección de sesgos en pronósticos meteorológicos numéricos.

Capítulo 5. Conclusiones y recomendaciones

Conclusiones

Basado en los resultados presentados y analizados en este estudio, se puede concluir que la implementación del Filtro de Kalman Adaptativo ha demostrado ser eficaz en la mejora de los pronósticos meteorológicos numéricos proporcionados por el sistema SREF. A través de la comparación entre los pronósticos brutos y corregidos, se ha observado una reducción significativa en el sesgo presente en las variables de temperatura máxima, temperatura mínima y precipitación.

En cuanto a los pronósticos del sistema SREF: para la variable de temperatura, se identificó que el sistema SREF tiende a subestimar la temperatura máxima en ambos trimestres analizados, así como la temperatura mínima durante los meses de MJJ. Por otro lado, se observa que el sistema tiende a sobreestimar la temperatura mínima durante el periodo DEF, así como la precipitación durante los periodos DEF y MJJ.

Al comparar los resultados de pronóstico con las observaciones, se evidencian diferencias significativas en las variables de temperatura máxima, temperatura mínima y precipitación. En general, se observa un sesgo promedio de -4.25 °C para la temperatura máxima y de 1.07 °C para la temperatura mínima. Durante los meses de DEF, el sesgo promedio de la temperatura es de 0.28 °C, mientras que durante los meses de MJJ es de -3.45 °C. En cuanto a la precipitación, se registra un sesgo promedio de 3.66 mm. Estos valores indican la diferencia presente entre los pronósticos y las observaciones.

A su vez, el sesgo absoluto promedio encontrado es de 5.27 °C para la temperatura máxima, 4.35 °C para la temperatura mínima, 4.35 °C para la variable de temperatura en general y 3.87 mm para la precipitación. Estos sesgos absolutos resaltan la magnitud de las diferencias entre los pronósticos y las observaciones.

Sin embargo, para los pronósticos corregidos al aplicar la corrección mediante el Filtro de

Kalman Adaptativo: se logra mejorar la precisión de los pronósticos. Los sesgos promedio disminuyen a -0.21 °C para la temperatura máxima, -0.14 °C para la temperatura mínima y -0.18 °C para la temperatura en general. En el caso de la precipitación, se obtiene un sesgo promedio de 0.11 mm.

El sesgo absoluto promedio también disminuye, mostrando una mejora en la precisión de las correcciones. Para la temperatura máxima, el sesgo absoluto promedio se reduce a 0.21 °C, mientras que para la temperatura mínima es de 0.14 °C y para la temperatura en general a 0.18 °C. Para la precipitación, el sesgo absoluto promedio es de 0.17 mm. Estos resultados indican una reducción en las diferencias entre las correcciones y las observaciones, lo que sugiere una mejora en la calidad de los pronósticos y una reducción de sesgo buena para las variables analizadas.

Además, los intervalos para la temperatura se reducen a un rango de $[-2$ a $2]$, excepto para la temperatura mínima durante los meses de MJJ, donde el rango es de $[-2$ a $0]$. Esto demuestra una mayor precisión en los pronósticos corregidos en comparación con los pronósticos en bruto, en donde los intervalos promedio para temperatura van desde $[-16.5$ a $13.5]$.

En el caso de la precipitación, si bien las mejoras no fueron tan significativas como en la temperatura, se evidenció una tendencia positiva al aplicar el método de corrección. La consideración de transformar los valores negativos en cero para evitar distorsiones en los resultados pudo haber afectado la magnitud de las mejoras, pero, aun así, se observa una reducción en el sesgo.

En relación a las otras métricas estadísticas adicionales planteadas en este trabajo, en el Anexo 2, donde se muestran los resultados del pronóstico numérico de SREF en comparación con las observaciones, se encontró que los intervalos de RMSE y MAE eran relativamente amplios, lo que indica una variabilidad considerable en los errores de los pronósticos.

Sin embargo, en el Anexo 3, que muestra los resultados de la corrección realizada por el Filtro de Kalman Adaptativo, se observó una notable mejora en la precisión de los

pronósticos. Los intervalos de RMSE y MAE fueron considerablemente más estrechos en comparación con los del Anexo 2, lo que indica una reducción significativa de las diferencias entre los valores observados y los pronosticados, y por consecuencia, de los errores. Además, se observó que un porcentaje considerable de estaciones presentaron valores más bajos de RMSE y MAE después de aplicar la corrección.

Para los coeficientes de correlación de Pearson (COE), se encontró una correlación positiva entre las variables observadas y las variables corregidas o pronosticadas, como se muestra en la Tabla 36 del Anexo 3. Sin embargo, se observó que la mejora en los valores de COE después de la aplicación del Filtro de Kalman Adaptativo no fue significativa en la mayoría de los casos. Se resalta que los coeficientes de correlación solo miden la relación lineal entre las variables y pueden no reflejar completamente la mejora en la reducción de errores entre los valores pronosticados y observados debido a la presencia de otros factores, como la dispersión de los datos o valores anómalos.

No obstante, en el ejemplo mostrado para la estación de Jiutepec, Mor., los resultados demuestran que al eliminar el primer día de análisis y calcular los coeficientes de correlación después de que el filtro haya tenido un impacto en la corrección, se logra una mejora significativa en los valores de COE. Esto respalda la eficacia de la corrección realizada por el Filtro de Kalman Adaptativo y sugiere una mayor consistencia entre los valores pronosticados y observados en términos de la relación lineal.

Adicionalmente se comprobó lo mencionado en, Pelosi, 2017., de que los valores iniciales de coeficientes de regresión $\bar{x}_{t|t-k\Delta t}$ y de la covarianza de su error $\bar{P}_{x,t|t-k\Delta t}$ convergen rápidamente a sus valores verdaderos.

En conclusión, la implementación del Filtro de Kalman Adaptativo ha permitido reducir de manera significativa las diferencias entre los pronósticos y las observaciones de temperatura máxima, temperatura mínima y precipitación. Los sesgos promedio y los sesgos absolutos promedio se han disminuido, lo que indica una menor diferencia entre los pronósticos corregidos por el Filtro de Kalman Adaptativo y las observaciones.

Los resultados obtenidos indican que el Filtro de Kalman Adaptativo es una herramienta prometedora para reducir el sesgo en los pronósticos meteorológicos numéricos y mejorar la calidad de los resultados. Este estudio ha demostrado que la implementación del filtro en el sistema SREF ha logrado una reducción significativa en el sesgo de los pronósticos de temperatura máxima, temperatura mínima y precipitación.

Se ha cumplido exitosamente el objetivo principal de disminuir las diferencias entre los resultados de las simulaciones numéricas y las observaciones meteorológicas en el territorio mexicano para estas variables. La hipótesis planteada también ha sido respaldada, ya que se ha observado una notable reducción en el sesgo entre los resultados de las simulaciones numéricas del modelo SREF y las observaciones puntuales de las variables meteorológicas mencionadas, al aplicar la técnica de corrección.

Al utilizar el filtro de Kalman Adaptativo, se logró corregir y ajustar los resultados de las simulaciones numéricas del sistema de ensamble SREF, lo que resultó en una disminución del sesgo en los pronósticos de las variables de interés. Esto indica una mayor precisión y confiabilidad de los pronósticos, lo cual es esencial para la toma de decisiones en diferentes áreas que dependen de la información meteorológica.

Estos hallazgos respaldan la relevancia del Filtro de Kalman Adaptativo como una herramienta de postprocesamiento para reducir el sesgo en los pronósticos meteorológicos numéricos y mejorar la calidad de los resultados. Sin embargo, se deben tener en cuenta las limitaciones y consideraciones en la aplicación del filtro, como ha sido mencionado para el caso de la precipitación.

Los resultados encontrados contribuyen al avance en el campo de la meteorología y proporcionan una base sólida para futuros estudios y mejoras en la calidad de los pronósticos numéricos del tiempo, a través de la reducción de sesgo mediante técnicas de postprocesamiento estadístico como lo es el Filtro de Kalman Adaptativo.

Recomendaciones

Aunque el objetivo principal de este trabajo no fue realizar una corrección de sesgo en tiempo real, se consideró la posibilidad de utilizar Estaciones Meteorológicas Automáticas (EMAs) para llevar a cabo dicha corrección en modo operativo. Sin embargo, no se pudo acceder a la información en tiempo real proveniente de estas estaciones debido a diversos desafíos. A pesar de esta limitación, se reconoce la importancia de este enfoque y se sugiere como una posible área de investigación futura para optimizar la corrección del sesgo en los pronósticos meteorológicos. Además, es relevante destacar que el uso de EMAs podría contribuir a mejorar la precisión de los pronósticos en tiempo real, especialmente cuando la atmósfera experimenta cambios drásticos en las condiciones atmosféricas.

La incorporación de EMAs en el proceso de pronóstico meteorológico proporcionaría datos actualizados sobre variables atmosféricas que no están disponibles a través de otras fuentes, como las estaciones meteorológicas convencionales. Estos datos podrían utilizarse posteriormente en el postprocesamiento a través del Filtro de Kalman Adaptativo de las salidas de los pronósticos meteorológicos numéricos. Los cambios repentinos en las condiciones atmosféricas representan un desafío para los pronosticadores, ya que los modelos meteorológicos a menudo tienen dificultades para capturar con precisión estos cambios y los fenómenos a pequeña escala. No obstante, al incorporar datos de EMAs en los modelos de pronóstico, se podría mejorar la capacidad de anticipar y adaptarse a estos cambios rápidos en tiempo real.

También, se aconseja realizar la aplicación de Filtro de Kalman Adaptativo a lo largo de varios años de información, con el fin de realizar un análisis del comportamiento de los coeficientes de regresión y demás variables calculadas por el algoritmo y lograr identificar patrones para el caso de cambios repentinos de las condiciones del tiempo atmosférico. Además, se sugiere aplicar la metodología generalizando para el caso múltiples predictores que expliquen el comportamiento del predictando.

Se recomienda llevar a cabo una investigación exhaustiva con el fin de identificar la metodología óptima para realizar la corrección del pronóstico numérico de la precipitación. En este contexto, se sugiere considerar el uso de técnicas como el Filtro de Kalman Extendido o el Filtro de Kalman Unscented, los cuales son aplicables a sistemas no lineales. Igualmente, se sugiere aplicar Filtro de Kalman Adaptativo, explotando toda la información del ensamble, como se ha empleado en otros estudios (Pelosi. 2017). Además, es relevante explorar la posibilidad de realizar la corrección de forma diferenciada por regiones geográficas o segmentando las estaciones según rangos altitudinales. Otra estrategia interesante sería evaluar la corrección utilizando datos de radar, ya que el Filtro de Kalman ha sido ampliamente utilizado para corregir el sesgo de la precipitación utilizando la reflectividad del radar.

Asimismo, es importante mencionar la opción de emplear la técnica conocida como Model Output Statistics (MOS), la cual consiste en identificar adecuadamente las variables que funcionen como predictores y utilizarlas en el proceso de corrección. La técnica MOS ha demostrado ser efectiva para mejorar los pronósticos meteorológicos, permitiendo ajustar y afinar las estimaciones basándose en datos observados y en modelos numéricos. Pero cuenta con ciertas limitaciones, como la necesidad de contar con una gran cantidad de datos históricos para desarrollar relaciones estadísticas efectivas.

Por todo lo anterior, se recomienda realizar un estudio riguroso y detallado para evaluar y comparar la eficacia de estas diversas metodologías de corrección en el contexto específico de la precipitación en México. Esta evaluación permitirá seleccionar la técnica más adecuada para obtener pronósticos precisos y confiables, lo cual contribuirá al avance de la predicción meteorológica y a la toma de decisiones basada en información precisa y oportuna.

Finalmente, se recomienda considerar el uso de técnicas de asimilación de datos, como WRFDA (Weather Research and Forecasting Data Assimilation) y WRF-DART (Weather Research and Forecasting Data Assimilation Research Testbed), para mejorar el rendimiento de los modelos de simulación numérica del tiempo. Estas técnicas permiten la integración de observaciones reales en los modelos numéricos, ajustando el estado inicial y mejorando la

precisión de los pronósticos. Estos programas, son ampliamente utilizados en la comunidad meteorológica, ofrecen métodos sofisticados, como la asimilación variacional y los Filtros de Kalman.

Referencias

- Beck, V. C., Yamasaki, Y. & Pereira, F. (2016). *El estado del arte de los métodos de asimilación de datos*. Universidad Federal de Pelotas. Facultad de Meteorología.
- Bjerknes, V. (1911). *Dynamical meteorological and hidrography*. New York: Carnegie Institute, Gibson Bros. 176p.
- Cassola, F., & Burlando, M. (2012). Wind speed and wind energy forecast through Kalman filtering of Numerical Weather Prediction model output. *Appl. Energy*, 99, 154–166, doi:10.1016/j.apenergy.2012.03.054.
- Castro, M., González, R., & Portela, A. (1991). *Parametrización de la capa límite atmosférica en los modelos numéricos de pronóstico meteorológico*. Dpto. Geofísica y Meteorología. Facultad de Física. Universidad Complutense 28040 Madrid. España.
- Caya, A., Sun, J., & Snyder, C. (2005). A comparison between the 4DVAR and the ensemble Kalman filter techniques for radar data assimilation. *Mon. Weather Rev.*, 133(11) 3081-3094.
- Cheng, W. Y. Y., & Steenburgh, W. J. (2007). Strengths and weaknesses of MOS, running-mean bias removal, and Kalman filter techniques for improving model forecasts over the western United States. *Wea. Forecasting*, 22, 1304–1318, doi:10.1175/2007WAF2006084.1.
- CICESE. (2013). *CLICOM*. Departamento de Oceanografía Física. División de Oceanología. Obtenido en: <http://clicom-mex.cicese.mx/>
- Coll-Hurtado, A. (1990). *Nuevo atlas nacional de México*. México: Universidad Nacional Autónoma de México, Instituto de Geografía. ISBN-13: 978-970-32-5047-9.
- COMET-MetEd. (2015a). *Operational Models Encyclopedia—SREF - Perturbation Method*. Obtenido en: <https://sites.google.com/ucar.edu/operational-models-encyclo/probabilistic-models/sref/ensemble-members>
- COMET-MetEd. (2015b). *Operational Models Encyclopedia—SREF - Perturbation Method*. Obtenido en: <https://sites.google.com/ucar.edu/operational-models-encyclo/probabilistic-models/sref/initial-conditions>
- COMET-MetEd. (2015c). *Operational Models Encyclopedia—SREF - Perturbation Method*. Obtenido en: <https://sites.google.com/ucar.edu/operational-models-encyclo/probabilistic-models/sref/vertical-resolution>
- COMET-MetEd. (2015d). *Operational Models Encyclopedia—SREF - Perturbation Method*. Obtenido en: <https://sites.google.com/ucar.edu/operational-models-encyclo/probabilistic-models/sref/grid-scale-resolution>
- COMET-MetEd. (2015e). *Operational Models Encyclopedia—SREF - Perturbation Method*. Obtenido en: <https://sites.google.com/ucar.edu/operational-models-encyclo/probabilistic-models/sref/ensemble-members>

models/sref/horizontal-resolution

Comisión Nacional del Agua. (2019). *Situación de los Recursos Hídricos (Capítulo 2)*. Obtenido en: <https://www.gob.mx/conagua/acciones-y-programas/situacion-de-los-recursos-hidricos>

Crochet, P. (2004). *Adaptive Kalman filtering of 2-metre temperature and 10-metre wind-speed forecasts in Iceland*. *Meteor. Appl.*, 11, 173–187, doi:10.1017/S1350482704001252.

Dillon, M. E., García, Y., & Collini, A. (s. f.). *Asimilación de Datos de un Modelo de Pronóstico Numérico del Tiempo en el Servicio Meteorológico Nacional*.

Delle Monache et al. (2006). *Ozone ensemble forecasts: 2. A Kalman filter predictor bias correction*. *Journal of Geophysical Research*. 111. 10.1029/2005JD006311.

Delle Monache, L., Nipen, T., Liu, Y., Roux, G., & Stull, R. (2011). *Kalman filter and analog schemes to postprocess numerical weather predictions*. *Mon. Wea. Rev.*, 139, 3554–3570, doi:10.1175/2011MWR3653.1

Dorado, J., & Ruíz, J. F. (2013). *Implementación de Filtros de Kalman como método de ajuste a los modelos de pronóstico (GFS) de temperaturas máxima y mínima para algunas ciudades de Colombia*. Grupo de Modelamiento de Tiempo y Clima. Subdirección de Meteorología – IDEAM.

ECMWF, 2015: *User guide to ECMWF forecast products v1.2*. ECMWF, 129 pp., http://www.ecmwf.int/sites/default/files/User_Guide_V1.2_20151123.pdf.

Evensen, G. (1994a). *Sequential data assimilation with a nonlinear quasigeostrophic model using Monte Carlo methods to forecast error statistics*. *J. Geophys. Res.*, 99(C5): 10143-10162.

Evensen, G. (1994b). *Inverse Methods and data assimilation in nonlinear ocean models*. *Physica D*, 77(1-3): 108-129.

Evensen, G. (2009). *Data Assimilation: The Ensemble Kalman Filter*. <https://link.springer.com/book/10.1007/978-3-642-03711-5>

Fischer, M. (2001). *Assimilation techniques (4): 4d-Var*. *Meteorological Training Course Lecture Series*, 2001.

Galanis, G., & Anadranistakis, M. (2002). *A one-dimensional Kalman filter for the correction of near surface temperature forecasts*. *Meteor. Appl.*, 9, 437–441, doi:10.1017/S1350482702004061.

Galanis, G., Louka, P., Katsafados, P., Pytharoulis, I., & Kallos, G. (2006). *Applications of Kalman filters based on non-linear functions to numerical weather predictions*. *Ann. Geophys.*, 24, 2451–2460, doi:10.5194/angeo-24-2451-2006.

Garduño, R. (1994). *EL VELEIDOSO CLIMA*. México: Fondo de Cultura Económica.

Glahn, H. R. & Lowry, D. A. (1972). *The use of Model Output Statistics (MOS) in Objective Weather Forecasting*. Techniques Development Laboratory. National Weather Service, NOAA.

Grupo de Interacción Océano Atmósfera. (2017). *WRF: Pronóstico Meteorológico*. Instituto de Ciencias de la Atmósfera y Cambio Climático, UNAM. Obtenido en: <http://grupo-ioa.atmosfera.unam.mx/pronosticos/index.php/meteorologia/inf-wrf>

Harms, D. E., Raman, S., & Madala, R. V. (1992). *An Examination of Four-Dimensional Data-Assimilation Techniques for Numerical Weather Prediction*. *Bull. Amer. Meteor. Soc.*, 73, No. 4, pp. 425–440.

Homleid, M. (1995). *Diurnal corrections of short-term surface temperature forecasts using the Kalman filter*. *Wea. Forecasting*, 10, 689–707, doi:10.1175/1520-0434(1995)010<0689:DCOSTS>2.0.CO;2.

Hunt, B. R., Kostelich, E. J., & Szunyogh, I. (2007). *Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter*. *Physica D: Nonlinear Phenomena*, 230, 112–126, doi:10.1016/j.physd.2006.11.008.

Jung, T. (2005) *Systematic errors of the atmospheric circulation in the ECMWF forecasting system*. *Quarterly Journal of Royal Meteorological Society*, 131, 1045–1073.

Kalnay, E. (2003). *Atmospheric modeling, data assimilation and predictability*. Cambridge University Press, Cambridge. 341p.

Libonati, R., Trigo, I., & DaCamara, C. C. (2008). *Correction of 2 m-temperature forecasts using Kalman filtering technique*. *Atmos. Res.*, 87, 183–197, doi:10.1016/j.atmosres.2007.08.006.

Lighezzolo, R. A. (2014) *Integración de modelos numéricos de predicción meteorológica en sistemas de alerta temprana a emergencias*. UNIVERSIDAD NACIONAL DE CÓRDOBA. Facultad de Matemática, Astronomía y Física –FaMAF

López de Llergo, R. (2003). *Principales rasgos geográficos de la República Mexicana*. *Investigaciones Geográficas*, (50), 7-23. Obtenido en: https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0188-46112003000100007

Meng, Z. & Zhang, F. (2008). *Tests of an ensemble Kalman filter for Mesoscale and regional-scale data assimilation. Part III: Comparison with 3DVAR in a real-data case study*. *Mon. Weather Rev.*, 136: 522-540.

Morss R., Demuth, J., Lazo, J. (2008). *Communicating Uncertainty in Weather Forecasts: A Survey of the U.S. Public*. *Weather and forecasting*; 23: 974-991.

Neslihanoglu, S., Ünal, E., & Yozgatligil, C. (2021). *Performance comparison of filtering methods on modelling and forecasting the total precipitation amount: a case study for Muğla in Turkey*. *Journal of Water and Climate Change*. 12. 10.2166/wcc.2021.332.

- Pelosi, A., Medina, H., Van den Bergh, J., Vannitsem, S., & Chirico, G. B. (2017). *Adaptative Kalman Filtering for Postprocessing Ensemble Numerical Weather Predictions*. 4837-4854. Obtenido en: <https://doi.org/10.1175/MWR-D-17-0084.1>
- Persson, A. (1991). *Kalman filtering—A new approach to adaptive statistical interpretation of numerical meteorological forecasts*. *Lectures and Papers Presented at the WMO Training Workshop on the Interpretation of NWP Products in Terms of Local Weather Phenomena and Their Verification*, H. R. Glahn, Ed., PSMP Report Series, Vol. 34, World Meteorological Organization, 27–32.
- Pielke, A. R. (1984). *Mesoscale Meteorological Modeling*. 978-0-08-092526-4. Obtenido en: <https://doi.org/10.1016/C2009-0-02981-X>
- Pinson, P. (2012). *Adaptive calibration of (u,v)-wind ensemble forecasts*. *Quart. J. Roy. Meteor. Soc.*, 138, 1273–1284, doi:10.1002/qj.1873.
- Python Software Foundation. (2023). *About Python*. Obtenido en: <https://www.python.org/about/>
- Quintas, I. (2000). *ERIC II: Documentación de la base de datos climatológica y del programa extractor*. Instituto Mexicano de Tecnología del Agua. 968-7417-40-4. Obtenido en: <http://repositorio.imta.mx/handle/20.500.12013/1635>
- Righetti, S. A., Hobouchian, M. P., & García, Y. (2019). *Corrección del BIAS del pronóstico de temperatura mínima y máxima del modelo GFS*. *Nota Técnica SMN 2017-56*. Servicio Meteorológico Nacional. Buenos Aires, Argentina.
- Sasaki, Y. (1958). *An objective analysis based on the variational method*. *Journal of Meteorological Society of Japan*, 36 (3): 77-88.
- Solera, R. Á. (2003). *El Filtro de Kalman*. Banco Central de Costa Rica División Económica Departamento de Investigaciones Económicas.
- Spyder IDE. (2023). *Spyder IDE*. Obtenido en: <https://www.spyder-ide.org/>
- Stathopoulos, C., Kaperoni, A., Galanis, G., & Kallos, G. (2013). *Wind power prediction based on numerical and statistical models*. *J. Wind Eng. Ind. Aerodyn.*, 112, 25–38, doi:10.1016/j.jweia.2012.09.004.
- Stauffer, D.R., Seaman, N.L. (1990). *Use of 4-D data assimilation in a limited area Mesoscale model. Part1: Experiments with synoptic scale data*. *Mon. Weather Rev.*, 118(6): 1250-1277.
- Talagrand, O. 1997. *Assimilation of observations: An introduction*. *Journal of the Meteorological Society of Japan*, 75, 191 – 205.
- Vannitsem, S. (2008). *Dynamical properties of MOS forecasts: Analysis of the ECMWF operational forecasting system*. *Wea. Pronóstico*, 23, 1032–1043, doi:10.1175/2008WAF2222126.1

- Velasquez, P., Messmer, M., & Raible, C. C. (2020). A new bias-correction method for precipitation over complex terrain suitable for different climate states: A case study using WRF (version 3.8.1). *Geoscientific Model Development*, 13(10), 5007-5027. <https://doi.org/10.5194/gmd-13-5007-2020>
- Warner T., Key, L., & Lario, A. (1989). Sensitivity of mesoscale model forecast skill to some initial data characteristics: data density, data position, analysis procedure and measurement error. *Mon. Wea. Rev.*, 117, o 6, 1281 – 1310.
- Welch, G., & Bishop, G. (2001). *An Introduction to the Kalman Filter*. Department of Computer Science University of North Carolina at Chapel Hill. Course 8, University of North Carolina at Chapel Hill. Department of Computer Science Chapel Hill.
- Xu, J., Xiao, Z., Lin, Z. et al. (2021). System bias correction of short-term hub-height wind forecasts using the Kalman filter. *Prot Control Mod Power Syst* 6, 37. Obtenido en: <https://doi.org/10.1186/s41601-021-00214-x>
- Zhang, M., Zhang, F., Huang, X.-Y, & Zhang, X. (2011). Intercomparison of an Ensemble Kalman Filter with Three- and Four- Dimension Variational Data Assimilation Methods in a Limited-Area Model over the Month of June 2003. *Mon. Weather Rev.*, 139: 566-572.

ANEXOS

Anexo 1: Acrónimos.

- 3DVAR: Three-dimensional variational assimilation
- 4DVAR: Four-dimensional variational assimilation
- AFWA: Air Force Weather Agency
- ARW: Advanced Research WPS
- WRF-ARW - Advanced Research WRF
- BIAS: Sesgo
- CAPS: Center for Analysis and Prediction Storms
- CLICOM: CLImate COMputing project
- COE: Coeficiente de correlación de Pearson
- DEF: Diciembre, enero y febrero
- DKF: Filtro de Kalman discreto
- EKF: Filtro de Kalman Extendido
- ECMWF: European Centre for Medium-Range Weather Forecast
- EMAs: Estaciones Meteorológicas Automáticas
- EnKF: Filtro de Kalman por Ensamblados
- ESRL: Earth System Research Laboratory
- FAA: Federal Aviation Administration
- GFS: Global Forecast System
- GDAS: Global Forecast System Data Assimilation System
- GrADS: Grid Analysis and Display System
- IDE: Integrated Development Environment
- KF: Filtro de Kalman
- LETKF: Local Ensemble Transform Kalman Filter
- MAE: Error absoluto medio
- MJJ: Mayo, junio y julio
- MOS: Model Output Statistics
- NAM: The North American Mesoscale Model
- NCAR: National Center for Atmospheric Research
- NCEP: National Center for Environmental Prediction

- NDAS: The North American Mesoscale Model Data Assimilation System
- NRL: Naval Research Laboratory
- NMM: Nonhydrostatic Mesoscale Model
- NMMB: NEMS Non-hydrostatic Multiscale Model on the B grid
- NOAA: National Oceanic and Atmospheric Administration
- NWP: Modelos de predicción numérica del tiempo
- RAP: Rapid Refresh
- RMSE: Raíz del error cuadrático medio
- SREF: Short-Range Ensemble Forecast
- SMN: Servicio Meteorológico Nacional
- WRF: Weather Research and Forecasting model

Anexo 2: RMSE, MAE y COE, pronóstico vs observación.

En este anexo se presentan los resultados obtenidos de las métricas estadísticas RMSE (raíz de error cuadrático medio) y MAE (error medio absoluto) calculadas entre el pronóstico del SREF y las observaciones de 1570 estaciones climáticas convencionales. Los resultados se muestran en forma de tablas, y a continuación se resumen los principales hallazgos de cada una de ellas:

Algo importante de mencionar es que valores de RMSE altos, indican que existe una mayor variabilidad entre los pronósticos y las observaciones. Lo que significa que, cuanto mayor sea el valor de RMSE, más grande será la diferencia entre los valores pronosticados y las observaciones. Lo que implicaría que los pronósticos pueden tener un nivel de precisión bajo.

Para el caso del MAE, representa el promedio de las diferencias absolutas entre pronósticos y los valores observados. Proporciona una medida de la magnitud promedio de los errores de pronóstico, independientemente de su dirección porque es absoluto. Un MAE bajo indica una mayor precisión en los pronósticos, ya que implica que los errores son relativamente pequeños en promedio.

Dicho lo anterior, un RMSE alto indica una mayor variabilidad y diferencias entre los pronósticos y las observaciones, lo cual sugiere una menor precisión en los pronósticos. Por otro lado, un MAE bajo indica una mayor precisión y menor magnitud promedio de los errores de pronóstico.

Tabla 20: RMSE para temperatura máxima en el trimestre DEF

En esta tabla, se observa que los intervalos de RMSE para la temperatura máxima, oscilando entre 0.00 y 22.0. La mayor cantidad de datos se concentra en el intervalo de RMSE entre 2.00 y 4.0, lo que representa el 37.51% de los datos.

Tabla 21: RMSE para temperatura máxima en el trimestre MJJ

Los intervalos de RMSE para la temperatura máxima en el trimestre MJJ varían entre 0 y 20.

La mayor cantidad de datos se encuentra en el intervalo de 2.0 a 4.0, lo que representa 440 estaciones de las 1570.

Tabla 22: RMSE para temperatura mínima en el trimestre DEF

En esta tabla, se observa que la mayor cantidad de datos se encuentra en el intervalo de temperatura mínima entre 2.0 y 4.0, lo que representa el 37.70 % de los datos. Los intervalos de RMSE van desde 0.00 hasta 20.00.

Tabla 23: RMSE para temperatura mínima en el trimestre MJJ

La mayor concentración de datos se encuentra en el intervalo de 2.0 a 4.0 para la temperatura mínima en el trimestre MJJ, con 513 de las 1570 estaciones. Los intervalos de RMSE llegan hasta valores de 12.

Tabla 24: MAE para temperatura máxima en el trimestre DEF

En esta tabla, se observa que 567 estaciones se encuentran en el intervalo de MAE entre 2 y 4 para la temperatura máxima en el trimestre DEF. Los intervalos de MAE van desde 0 hasta 20.

Tabla 25: MAE para temperatura máxima en el trimestre MJJ

Nuevamente, la mayor cantidad de datos se encuentra en el intervalo de MAE entre 2 y 4 para la temperatura máxima en el trimestre MJJ, con 387 estaciones dentro de ese intervalo. Los intervalos de MAE van desde 0 hasta 20.

Tabla 26: MAE para temperatura mínima en el trimestre DEF

El 38.66% de las estaciones presenta un MAE entre 2 y 4 para la temperatura mínima en el trimestre DEF. Los intervalos de MAE van desde 0 hasta 18.

Tabla 27: MAE para temperatura mínima en el trimestre MJJ

En esta tabla, el 43.4% de los datos tiene un MAE entre 2 y 4 para la temperatura mínima en el trimestre MJJ. Los intervalos de MAE van desde 0 hasta 12.

Estas tablas proporcionan información detallada sobre los resultados de los análisis estadísticos realizados, permitiendo una comprensión más precisa de la precisión de los pronósticos del SREF en comparación con las observaciones de las estaciones climáticas convencionales. Los intervalos de RMSE y MAE muestran la variabilidad en los errores de pronóstico y cómo se distribuyen los datos dentro de cada intervalo.

Estos resultados son relevantes para evaluar la calidad y confiabilidad de los pronósticos del SREF en relación con las condiciones observadas.

Tabla 20: Intervalos de RMSE para la temperatura máxima en el trimestre de diciembre, enero y febrero 2020-2021, pronóstico vs observación.

INTERVALO	NUM. DATOS	%
[0.0, 2.0]	29	1.84713376
[2.0, 4.0]	589	37.5159236
[4.0, 6.0]	421	26.8152866
[6.0, 8.0]	272	17.3248408
[8.0, 10.0]	175	11.1464968
[10.0, 12.0]	45	2.86624204
[12.0, 14.0]	25	1.59235669
[14.0, 16.0]	12	0.76433121
[16.0, 18.0]	1	0.06369427
[18.0, 20.0]	0	0
[20.0, 22.0]	1	0.06369427

Tabla 21: Intervalos de RMSE para la temperatura máxima en el trimestre de mayo, junio y julio 2020, pronóstico vs observación

INTERVALO	NUM. DATOS	%
[0.0, 2.0]	104	6.62420382
[2.0, 4.0]	440	28.0254777
[4.0, 6.0]	188	11.9745223
[6.0, 8.0]	231	14.7133758
[8.0, 10.0]	190	12.1019108
[10.0, 12.0]	179	11.4012739
[12.0, 14.0]	127	8.08917197
[14.0, 16.0]	90	5.73248408
[16.0, 18.0]	20	1.27388535

[18.0, 20.0]	1	0.06369427
--------------	---	------------

Tabla 22: . Intervalos de RMSE para la temperatura mínima en el trimestre de diciembre, enero y febrero 2020-2021, pronóstico vs observación.

INTERVALO	NUM. DATOS	%
[0.0, 2.0]	23	1.46496815
[2.0, 4.0]	592	37.7070064
[4.0, 6.0]	347	22.1019108
[6.0, 8.0]	208	13.2484076
[8.0, 10.0]	212	13.5031847
[10.0, 12.0]	95	6.05095541
[12.0, 14.0]	71	4.52229299
[14.0, 16.0]	16	1.01910828
[16.0, 18.0]	5	0.31847134
[18.0, 20.0]	1	0.06369427

Tabla 23: Intervalos de RMSE para la temperatura mínima en el trimestre de mayo, junio y julio 2020, pronóstico vs observación.

INTERVALO	NUM. DATOS	%
[0.0, 2.0]	170	10.8280255
[2.0, 4.0]	513	32.6751592
[4.0, 6.0]	421	26.8152866
[6.0, 8.0]	289	18.4076433
[8.0, 10.0]	155	9.87261146
[10.0, 12.0]	22	1.40127389

Tabla 24: Intervalos de MAE para la temperatura máxima en el trimestre de diciembre, enero y febrero 2020-2021, pronóstico vs observación

INTERVALO	NUM. DATOS	%
[0.0, 2.0]	214	13.6305732
[2.0, 4.0]	567	36.1146497
[4.0, 6.0]	357	22.7388535
[6.0, 8.0]	238	15.1592357
[8.0, 10.0]	128	8.15286624
[10.0, 12.0]	37	2.3566879
[12.0, 14.0]	21	1.33757962
[14.0, 16.0]	7	0.44585987
[16.0, 18.0]	0	0
[18.0, 20.0]	1	0.06369427

Tabla 25: Intervalos de MAE para la temperatura máxima en el trimestre de mayo, junio y julio 2020, pronóstico vs observación.

INTERVALO	NUM. DATOS	%
[0.0, 2.0]	216	13.7579618
[2.0, 4.0]	387	24.6496815
[4.0, 6.0]	171	10.8917197
[6.0, 8.0]	220	14.0127389
[8.0, 10.0]	197	12.5477707
[10.0, 12.0]	156	9.93630573
[12.0, 14.0]	123	7.8343949
[14.0, 16.0]	82	5.22292994
[16.0, 18.0]	17	1.08280255
[18.0, 20.0]	1	0.06369427

Tabla 26: Intervalos de MAE para la temperatura mínima en el trimestre de diciembre, enero y febrero 2020-2021, pronóstico vs observación.

INTERVALO	NUM. DATOS	%
[0.0, 2.0]	153	9.74522293
[2.0, 4.0]	607	38.6624204
[4.0, 6.0]	250	15.9235669
[6.0, 8.0]	182	11.5923567
[8.0, 10.0]	212	13.5031847
[10.0, 12.0]	86	5.47770701
[12.0, 14.0]	62	3.94904459
[14.0, 16.0]	12	0.76433121
[16.0, 18.0]	6	0.38216561

Tabla 27: Intervalos de MAE para la temperatura mínima en el trimestre de mayo, junio y julio 2020, pronóstico vs observación.

INTERVALO	NUM. DATOS	%
[0.0, 2.0]	312	19.8726115
[2.0, 4.0]	434	27.6433121
[4.0, 6.0]	402	25.6050955
[6.0, 8.0]	260	16.5605096
[8.0, 10.0]	143	9.10828025
[10.0, 12.0]	19	1.21019108

Anexo 3: RMSE, MAE y COE, corrección vs observación.

En este Anexo 3, se presentan los resultados obtenidos al evaluar las métricas estadísticas en relación con la corrección realizada por el Filtro de Kalman Adaptativo versus la observación. Se han calculado el RMSE (Error Cuadrático Medio), el MAE (Error Absoluto Medio) y el COE (Error de Correlación) para analizar la precisión de las correcciones realizadas en comparación con los datos observados. A continuación, se detalla el resumen de cada una de las tablas, que proporcionan información valiosa sobre el desempeño de las correcciones.

Tabla 28: RMSE para temperatura máxima DEF

En esta tabla se presentan los intervalos de RMSE para la temperatura máxima en el trimestre DEF. Se observa que los intervalos abarcan valores desde 0 a 6. La mayor proporción de datos se encuentra en el intervalo de RMSE entre 2 y 4, que engloba poco más del 84.14 % de los datos.

Tabla 29: RMSE para temperatura máxima MJJ

La tabla muestra los intervalos de RMSE para la temperatura máxima en el trimestre MJJ. Destaca que 1501 estaciones se encuentran en un intervalo de RMSE entre 2.0 y 4.0. Los intervalos de RMSE varían solo de 0 a 4.0.

Tabla 30: RMSE para temperatura mínima DEF

En esta tabla se presentan los intervalos de RMSE para la temperatura mínima en el trimestre DEF. Se observa que el 75.28% de las estaciones se encuentran en un intervalo de RMSE entre 0.0 y 2.0, y 382 estaciones se ubican en un intervalo de RMSE entre 2.0 y 4.0. Los intervalos de RMSE abarcan valores que van desde 0 a 6.

Tabla 31: RMSE para temperatura mínima MJJ

En la tabla se muestran los intervalos de RMSE para la temperatura mínima en el trimestre MJJ. Se encuentra que 1551 estaciones de las 1570 tienen un RMSE en el intervalo de 0 a 2. Los intervalos de RMSE varían desde 0 hasta 4.0.

Tabla 32: MAE para temperatura máxima DEF

En esta tabla se analiza el MAE (Error Absoluto Medio) para la temperatura máxima en el trimestre DEF. Los intervalos de MAE abarcan valores que van desde 0 hasta 6. Se observa que el 70.95% de las estaciones tienen un MAE en el rango de 0 a 2.

Tabla 33: MAE para temperatura máxima MJJ

La tabla muestra el MAE para la temperatura máxima en el trimestre MJJ. Destaca que el 98.89% de los datos presentan un MAE entre 0 y 2. El valor máximo del MAE es de 4.

Tabla 34: MAE para temperatura mínima DEF

En esta tabla se presenta el MAE para la temperatura mínima en el trimestre MJJ. Se observa que el 91.01 % de las estaciones presentan un MAE entre 0 y 2.

Tabla 35: MAE para temperatura mínima MJJ

En esta tabla se presenta el MAE para la temperatura mínima en el trimestre MJJ. Se observa que el 99.93% de las estaciones presentan un MAE entre 0 y 2.

Estos resultados proporcionan información detallada sobre la efectividad de las correcciones realizadas al aplicar el método de corrección de sesgo en comparación con las observaciones puntuales de las estaciones climáticas convencionales.

Tabla 28: Intervalos de RMSE para la temperatura máxima en el trimestre de diciembre, enero y febrero 2020-2021, corrección vs observación.

INTERVALO	NUM. DATOS	%
[0.0, 2.0]	176	11.2101911
[2.0, 4.0]	1321	84.1401274
[4.0, 6.0]	73	4.64968153

Tabla 29: Intervalos de RMSE para la temperatura máxima en el trimestre de mayo, junio y julio 2020, corrección vs observación.

INTERVALO	NUM. DATOS	%
[0.0, 2.0]	69	4.39490446
[2.0, 4.0]	1501	95.6050955

Tabla 30: Intervalos de RMSE para la temperatura mínima en el trimestre de diciembre, enero y febrero 2020-2021, corrección vs observación.

INTERVALO	NUM. DATOS	%
[0.0, 2.0]	1182	75.2866242
[2.0, 4.0]	382	24.3312102
[4.0, 6.0]	6	0.38216561

Tabla 31: Intervalos de RMSE para la temperatura mínima en el trimestre de mayo, junio y julio 2020, corrección vs observación.

INTERVALO	NUM. DATOS	%
[0.0, 2.0]	1551	98.7898089
[2.0, 4.0]	19	1.21019108

Tabla 32: Intervalos de MAE para la temperatura máxima en el trimestre de diciembre, enero y febrero 2020-2021, corrección vs observación.

INTERVALO	NUM. DATOS	%
[0.0, 2.0]	1114	70.955414
[2.0, 4.0]	455	28.9808917
[4.0, 6.0]	1	0.06369427

Tabla 33: Intervalos de MAE para la temperatura máxima en el trimestre de mayo, junio y julio 2020, corrección vs observación.

INTERVALO	NUM. DATOS	%
[0.0, 2.0]	1554	98.9808917
[2.0, 4.0]	16	1.01910828

Tabla 34: Intervalos de MAE para la temperatura mínima en el trimestre de diciembre, enero y febrero 2020-2021, corrección vs observación.

INTERVALO	NUM. DATOS	%
[0.0, 2.0]	1429	91.0191083
[2.0, 4.0]	139	8.85350318
[4.0, 6.0]	2	0.12738854

Tabla 35: Intervalos de MAE para la temperatura mínima en el trimestre de mayo, junio y julio 2020, corrección vs observación.

INTERVALO	NUM. DATOS	%
[0.0, 2.0]	1569	99.9363057
[2.0, 4.0]	1	0.06369427

Finalmente, la Tabla 36 muestra los resultados obtenidos del coeficiente de correlación de Pearson entre el pronóstico vs observación (COE SREF VS OBS) y el pronóstico del sistema SREF y los valores observados (COE SREF VS OBS). EL COE es una medida de la relación lineal entre dos variables, y varían de -1 a 1, donde 1 indica una correlación positiva perfecta, 0 indica ausencia de correlación y -1 indica una correlación negativa perfecta.

Para la variable de precipitación - DEF, el COE FK VS OBS es de 0.19, mientras que para COE SREF VS OBS es de 0.29. En cuanto, a la precipitación – MJJ, el COE FK VS OBS encontrado fue de 0.49 y el COE SREF VS OBS de 0.27. Estos coeficientes indican una correlación positiva débil en ambos casos, pero la correlación es un poco mayor en el caso de FK VS OBS para los meses MMJ y, para los meses DEF la correlación es mayor entre el SREF VS OBS.

Respecto a la temperatura máxima – DEF y MJJ, se obtienen coeficientes de correlación de 0.66 y 0.41 respectivamente para COE FK VS OBS. Para SREF VS OBS, los coeficientes de correlación son 0.62 y 0.34 respectivamente. Estos coeficientes indican una correlación positiva moderada en ambos casos, pero no hay una mejora significativa entre el COE promedio encontrado para estas variables y periodos.

En cuanto a la temperatura mínima - DEF y MJJ, se obtienen coeficientes de correlación de 0.73 y 0.78 respectivamente para FK VS OBS. En el caso de SREF VS OBS, el COE es de 0.49 y 0.34 respectivamente. Estos coeficientes indican una correlación positiva moderada a fuerte para el caso de FK VS OBS, pero nuevamente no hay una mejora significativa entre FK VS OBS y SREF VS OBS.

Tabla 36: Coeficiente de correlación de Pearson promedio encontrado para cada caso de estudio.

Variable - periodo	COE FK VS OBS	COE SREF VS OBS
Precipitación - MJJ	0.49	0.27
Precipitación - DEF	0.19	0.29
Temperatura máxima - DEF	0.66	0.62
Temperatura máxima - MJJ	0.41	0.34
Temperatura mínima - DEF	0.73	0.49
Temperatura mínima - MJJ	0.78	0.34

Los resultados obtenidos en la Tabla 36, revelan la existencia de una correlación positiva entre las variables observadas y las variables corregidas o pronosticadas. Sin embargo, se observa que, en la mayoría de los casos, la mejora no es significativa después de aplicar el Filtro de Kalman Adaptativo. Es importante tener en cuenta que los coeficientes de correlación solo evalúan la relación lineal entre las variables, y pueden no reflejar de manera precisa la reducción de las diferencias entre los valores pronosticados y observados debido a la influencia de otros factores, como la dispersión de los datos o valores atípicos.

Para abordar este aspecto, se decidió calcular los coeficientes de correlación después de eliminar el primer día de análisis. De esta manera, se aseguró contar únicamente con los días en los que los coeficientes de correlación fueron actualizados por el filtro y tuvieron un impacto en la corrección. Esto permitió reducir la anomalía observada en el primer día de análisis y eliminar una fuente potencial de error. Como ejemplo, se muestran gráficas de dispersión, entre la corrección vs observación (izquierda) y pronóstico vs observación (derecha) para la estación de Jiutepec, perteneciente al estado de Morelos

En las Figuras 38, 39, 40 y 41, para el caso de temperatura, se puede apreciar una clara mejoría en los valores de los coeficientes de correlación después de aplicar el método de corrección. Estos resultados respaldan la eficacia de la corrección realizada y sugieren una mayor concordancia entre los valores pronosticados y observados en términos de la relación lineal. A continuación, se muestra una descripción general de las figuras:

En la Figura 38, para temperatura máxima – DEF, se logra observar que el COE pasó de 0.72, a 0.78 al aplicar el Filtro de Kalman para la estación de Jiutepec, lo que muestra una mejora, aunque no tan significativa.

En la Figura 39, se muestra una corrección significativa para la variable de temperatura máxima – MJJ, teniendo un COE de 0.71 entre la corrección y la observación, a comparación de los 0.11 encontrados entre el sistema SREF y los valores observados.

En la Figura 40, representa la temperatura mínima – DEF para la estación de Jiutepec, se logra apreciar que el COE paso de ser 0.51 a 0.78 al implementar la técnica de corrección.

En la Figura 41, para la temperatura mínima – MJJ, el pronóstico del sistema SREF muestra un COE negativo muy bajo, de -0.18, para este caso específico, se observa una mejoría muy significativa al presentarse un COE después de aplicar el Filtro de Kalman Adaptativo de 0.91.

En la Figura 42, para la precipitación - DEF, los COE encontrados con la corrección y la observación son muy bajos, con valores de 0.06 y 0.12, respectivamente.

En Figura 43, para la precipitación – MJJ, se muestra una mejoría considerable, al pasar de un COE de 0.23 a uno de 0.51.

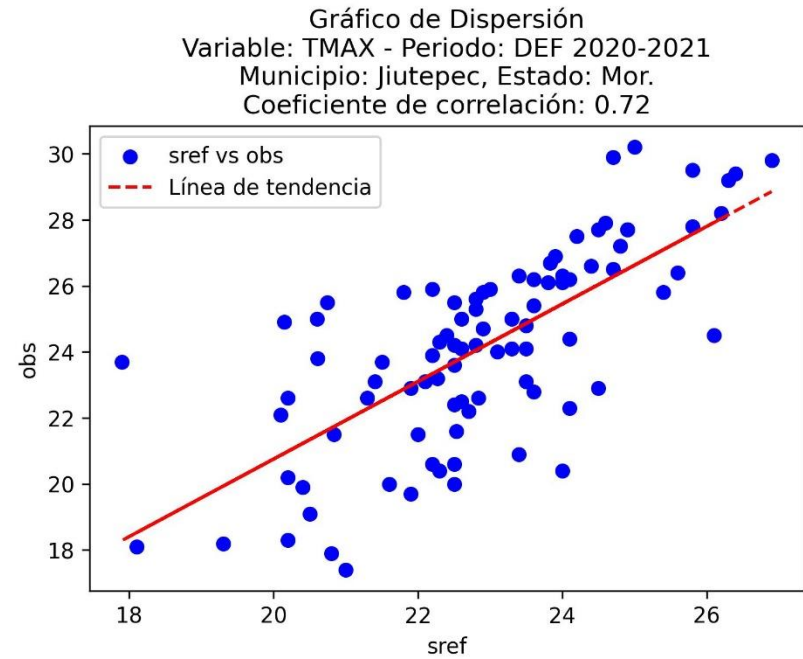
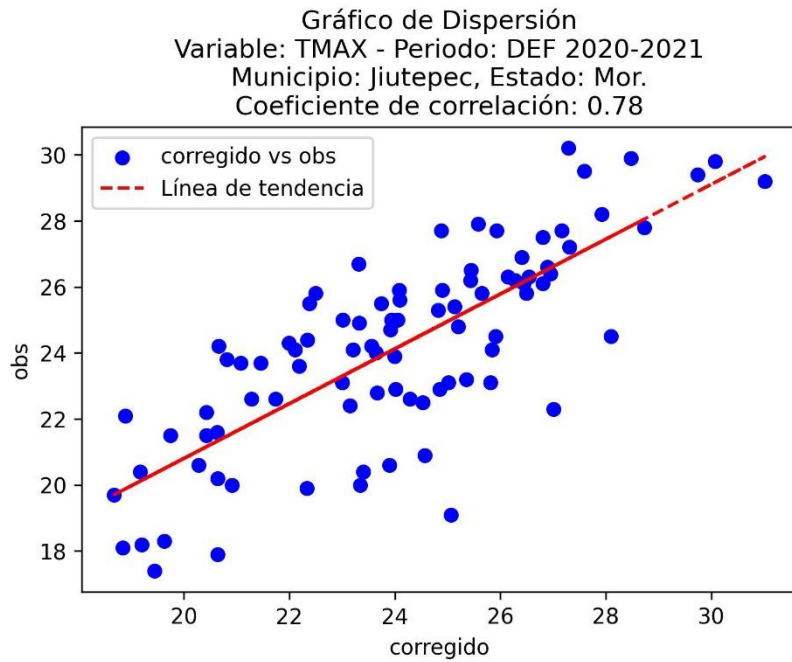


Figura 38: Comparación de la dispersión entre la temperatura máxima corregida (izquierda) y la pronosticada (derecha), contra las observaciones para la estación de Jiutepec, Mor., periodo DEF.

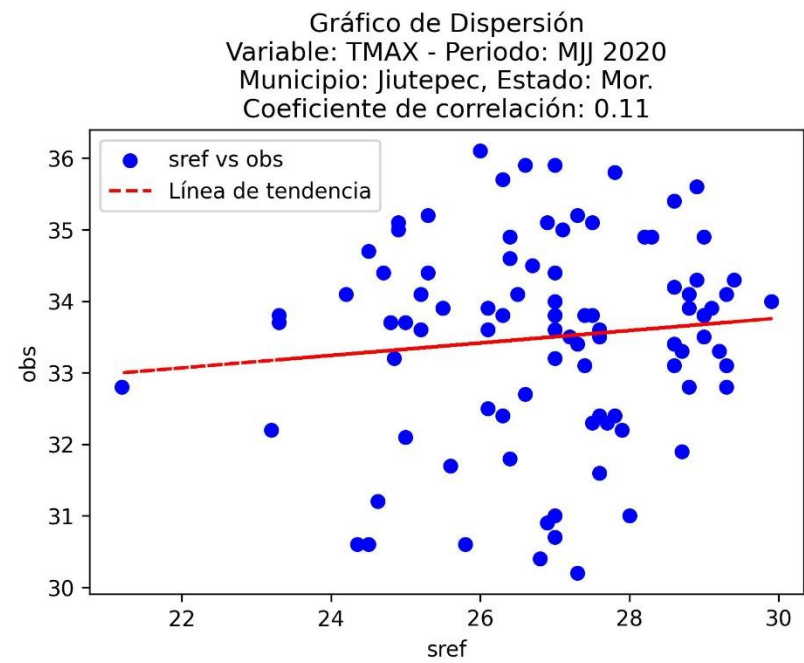
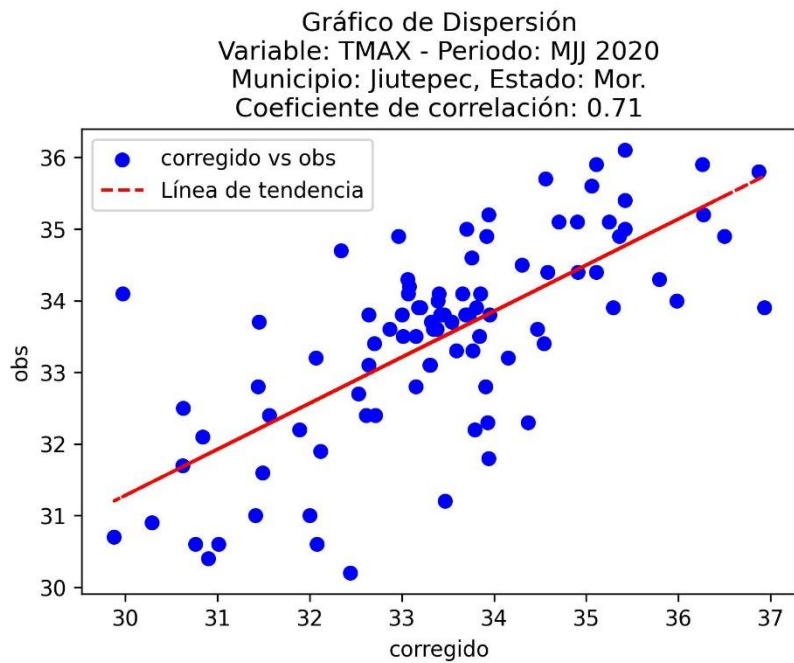


Figura 39: Comparación de la dispersión entre la temperatura máxima corregida (izquierda) y la pronosticada (derecha), contra las observaciones para la estación de Jiutepec, Mor., periodo MJJ.

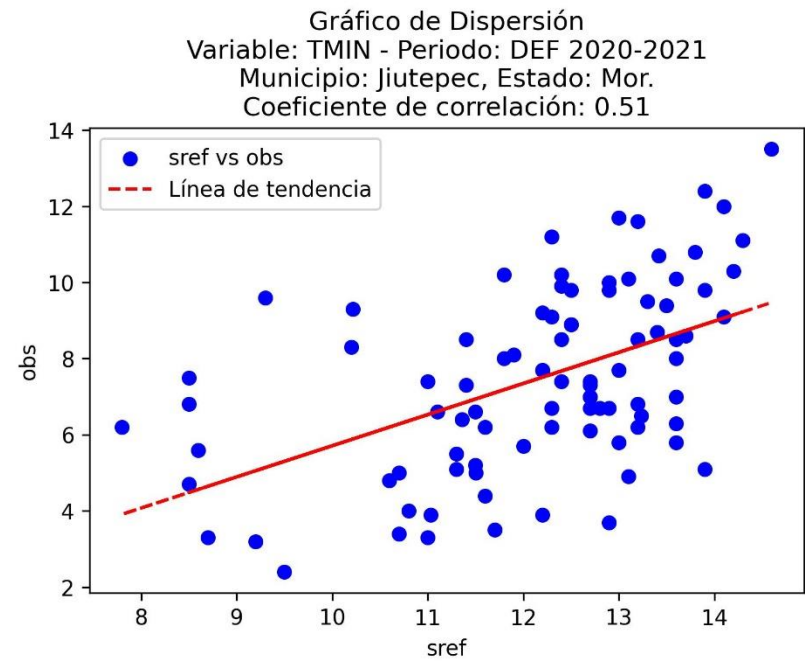
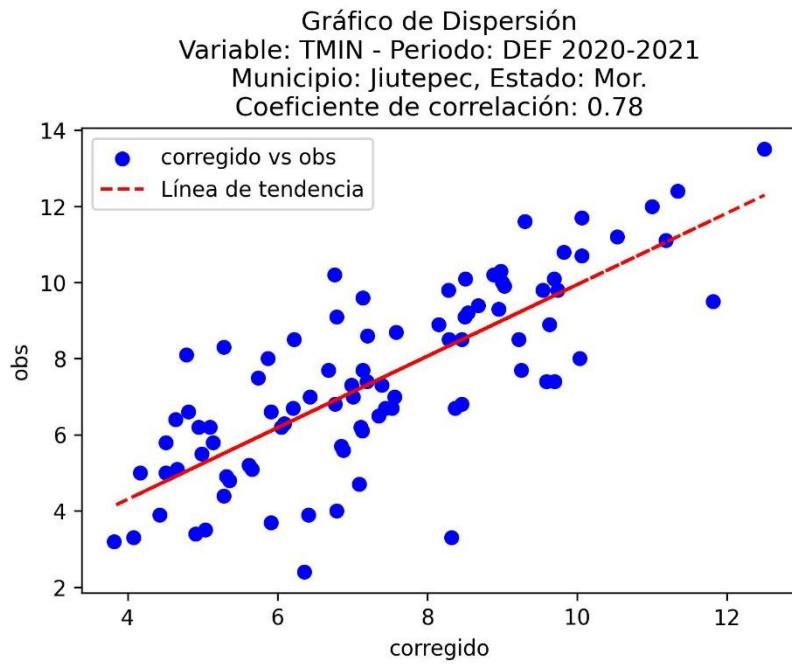


Figura 40: Comparación de la dispersión entre la temperatura mínima corregida (izquierda) y la pronosticada (derecha), contra las observaciones para la estación de Jiutepec, Mor., periodo DEF.

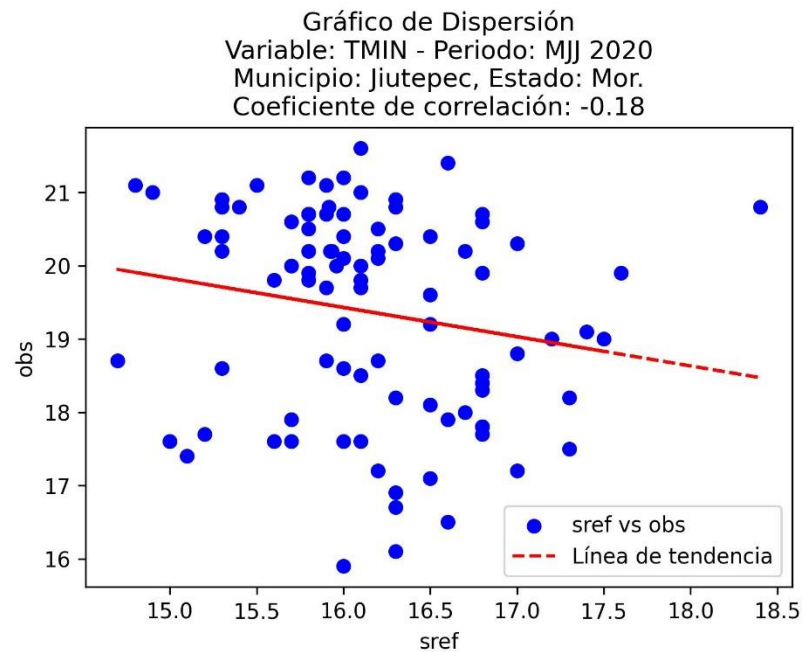
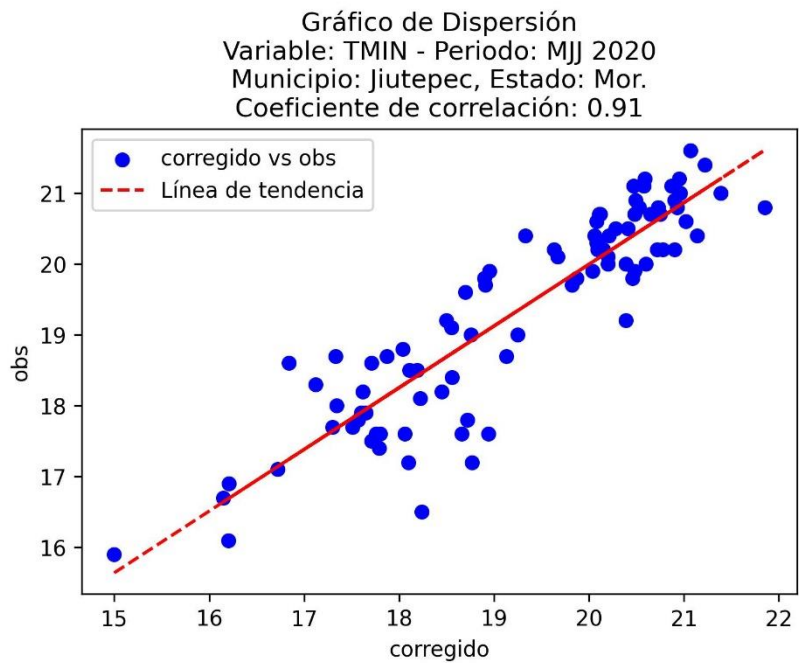


Figura 41: Comparación de la dispersión entre la temperatura mínima corregida (izquierda) y la pronosticada (derecha), contra las observaciones para la estación de Jiutepec, Mor., periodo MJJ.

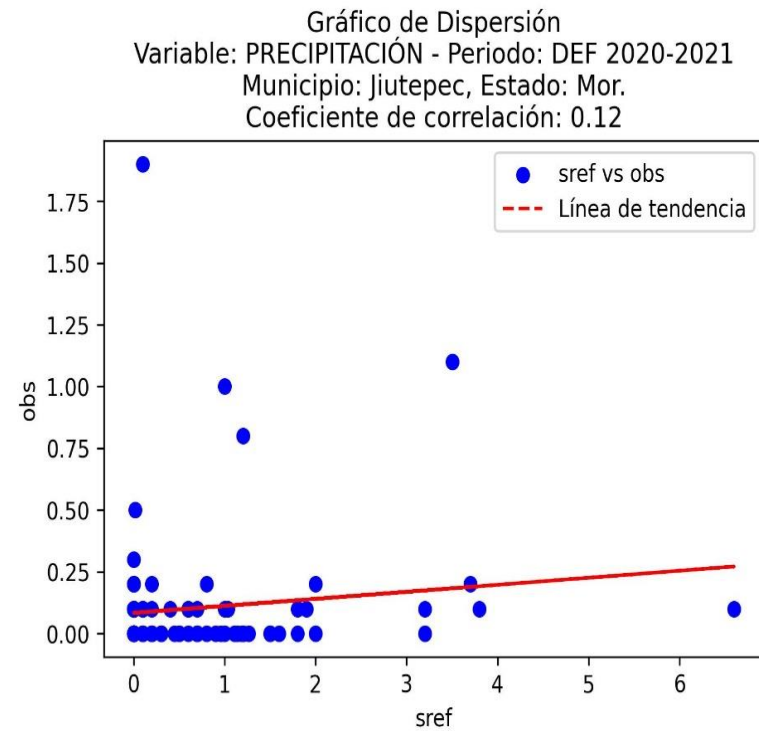
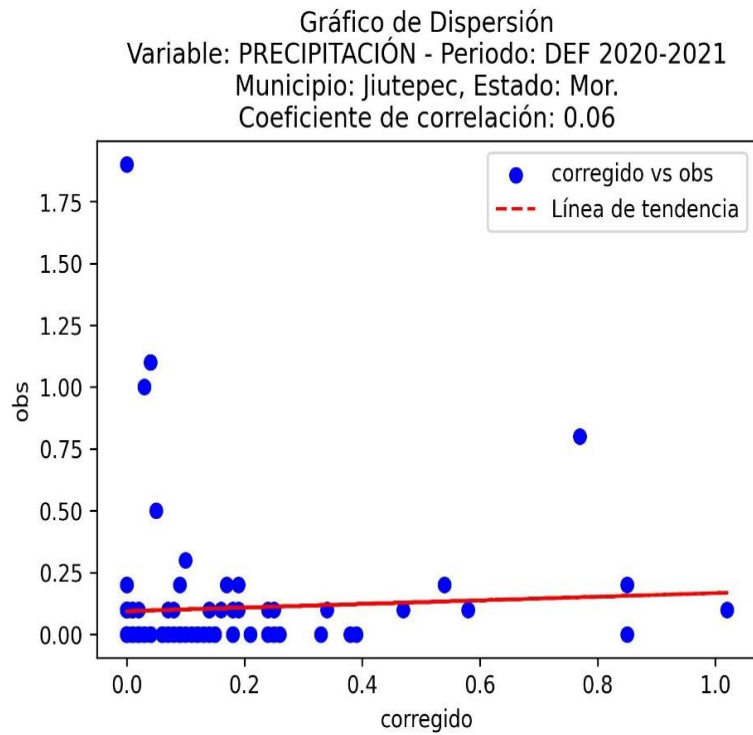


Figura 42: Comparación de la dispersión entre la precipitación corregida (izquierda) y la pronosticada (derecha), contra las observaciones para la estación de Jiutepec, Mor., periodo DEF

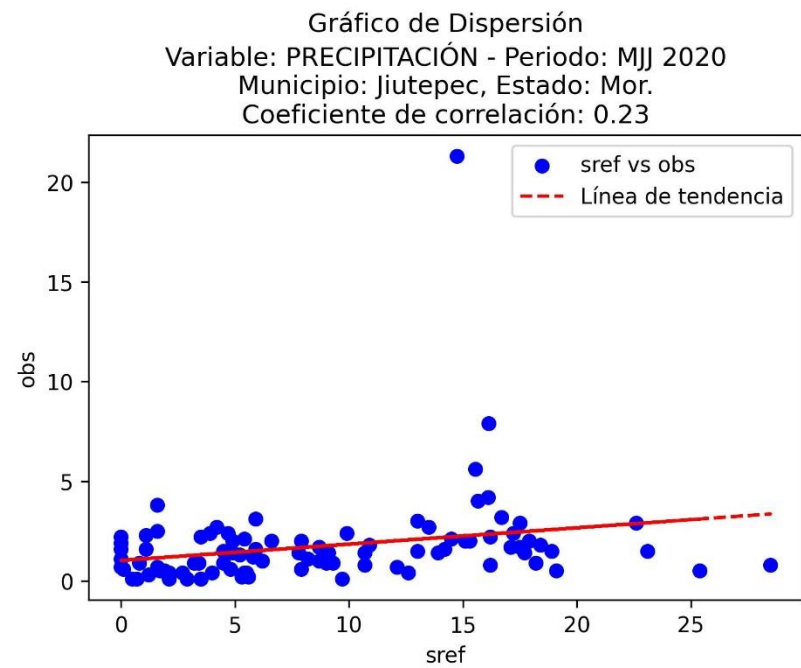
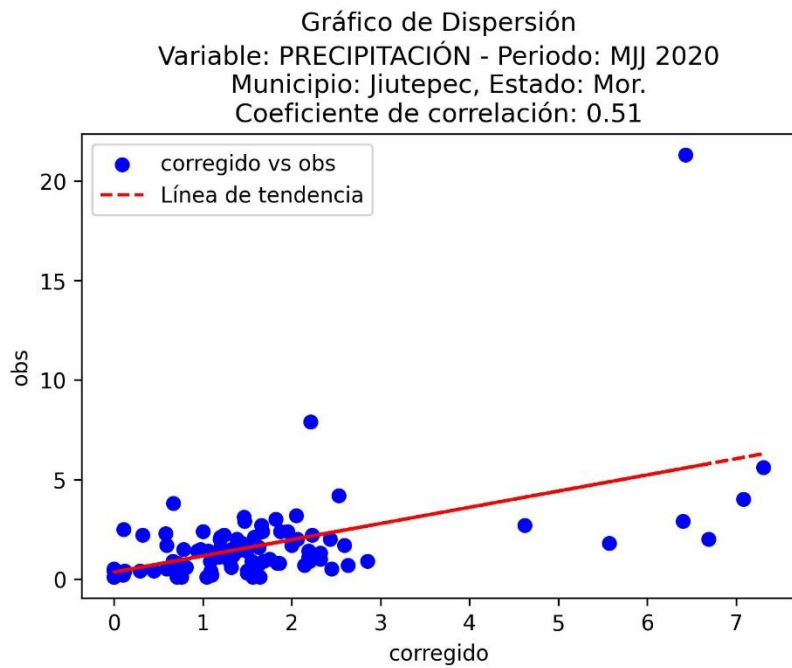


Figura 43: Comparación de la dispersión entre la precipitación corregida (izquierda) y la pronosticada (derecha), contra las observaciones para la estación de Jiutepec, Mor., periodo MJJ.

Análisis de los resultados mostrados en el Anexo 2 y 3

En el Anexo 2, muestra el RMSE y MAE del pronóstico numérico de SREF en comparación con la observación registrada, se observa que los intervalos de RMSE y MAE son bastante amplios en general. Esto indica que los pronósticos del sistema SREF presentan una variabilidad considerable en sus errores, con valores muy altos en la mayoría de los casos.

En el Anexo 3, que muestra el RMSE y MAE de la corrección realizada por el Filtro de Kalman Adaptativo, se aprecia que los intervalos de RMSE y MAE son mucho más estrechos en comparación con los del Anexo 2. Esto indica que la aplicación del Filtro de Kalman ha contribuido a reducir los errores en los pronósticos, mejorando la precisión en general de la estimación del sistema SREF. Además, se observa que un porcentaje significativo de estaciones se concentra en valores de RMSE y MAE más bajos, en comparación con lo encontrado en el Anexo 3.

Los resultados mostrados en final del Anexo 3, en la Tabla 36, de los coeficientes de correlación de Pearson (COE) indican que hay una correlación positiva entre las variables observadas y las variables corregidas o pronosticadas, sin embargo, después de aplicar el Filtro de Kalman Adaptativo la mejora no es significativa en la mayoría de los casos. Es importante tener en cuenta que los coeficientes de correlación solo miden la relación lineal entre las variables. El COE puede no reflejar la mejora en la reducción de los valores pronosticados y observados debido a la presencia de otros factores, como la dispersión de los datos o valores anómalos.

Sin embargo, al calcular el COE al excluir el primer día de análisis y considerar solo los días en los que los coeficientes de correlación fueron actualizados por el filtro y tuvieron un impacto en la corrección, se observó una notable mejora en los valores de COE. Esto se evidencia en las Figuras 38, 39, 40 y 41, donde se muestra una clara mejora en los valores de COE para las variables de temperatura después de aplicar el método de corrección.