

تطبيقات الذكاء الاصطناعي والاستشعار عن بُعد لتحليل جودة الهواء في الدمام والجبيل بالمملكة العربية السعودية جمعه بنت إبراهيم بن أحمد عین

قسم الجغرافيا، كلية العلوم الإنسانية والاجتماعية، جامعة الملك سعود، الرياض، المملكة العربية السعودية

Jain@ksu.edu.sa

المستخلص. تواجه مدينتا الدمام والجبيل تحديات بيئية وصحية نتيجة ارتفاع تلوث الهواء الناجم عن الأنشطة الصناعية وحركة المرور. هدفت الدراسة إلى تقييم جودة الهواء في المدينتين للفترة ٢٠١٩-٢٠٢٤ بالاعتماد على بيانات الأقمار الصناعية (Sentinel-5P وMODIS) ونماذج الذكاء الاصطناعي. أظهرت النتائج أن الجبيل سجلت تركيزات أعلى من الجسيمات الدقيقة ($PM_{2.5}$) بمتوسط ٤٥,٢٨ ميكروغرام/م³ مقارنة بالدمام (٤٢,٠٤ ميكروغرام/م³)، بالإضافة إلى ارتفاع مستويات NO_2 و SO_2 في الجبيل نتيجة التأثير الصناعي، بينما لم يلاحظ فرق كبير في تركيزات CO بين المدينتين. استخدم نموذج XGBoost للتنبؤ بجودة الهواء، وحقق دقة عالية (R^2 بين ٠,٩٦ و ٠,٩٨)، مع ملاحظة أن الأخطاء التنبؤية كانت أكبر في الجبيل. أكدت الدراسة أن الانبعاثات الصناعية وحركة المرور هما المصدران الرئيسيان للتلوث، مع تحسن نسبي لبعض الملوثات بفضل السياسات البيئية. توصي الدراسة بتعزيز كفاءة المركبات، زيادة المساحات الخضراء، وتطوير سياسات بيئية أكثر صرامة للحد من ارتفاع NO_2 و O_3 وتحقيق بيئة أكثر استدامة.

كلمات مفتاحية: تلوث الهواء، التنبؤات المستقبلية، التغيرات المكانية، المناطق الصناعية، النمذجة التنبؤية، الجغرافيا التطبيقية..

١- المقدمة

يُعد تلوث الهواء من القضايا البيئية البارزة ذات التأثير المباشر على صحة الإنسان والأنظمة البيئية، حيث تتسبب الانبعاثات الغازية والجسيمات الدقيقة في تدهور جودة الهواء، مما يؤدي إلى مشكلات صحية مزمنة مثل أمراض الجهاز التنفسي وأمراض القلب والأوعية الدموية (Ng et al., 2019). ووفقاً لمنظمة الصحة العالمية، يُسهم تلوث

الهواء في وفاة ملايين الأشخاص سنوياً، نتيجة التعرض المطول لمستويات مرتفعة من ملوثات مثل ثاني أكسيد النيتروجين (NO_2)، وثاني أكسيد الكبريت (SO_2)، وأول أكسيد الكربون (CO)، والجسيمات الدقيقة ($\text{PM}_{2.5}$)، والأوزون (O_3) (Banerjee et al., 2024). وتعتمد شدة هذا التلوث على عدة عوامل تشمل الأنشطة الصناعية، والانبعاثات الناتجة عن وسائل النقل، والظروف المناخية، فضلاً عن العوامل الجغرافية والطبوغرافية (Al-Taai & Mohammed al-Dulaimi, 2022).

تواجه المدن السعودية، وخاصة المناطق الصناعية الكبرى مثل الدمام والجبيل، تحديات متزايدة فيما يتعلق بجودة الهواء، نظراً للأنشطة الصناعية واسعة النطاق والتوسع الحضري السريع (AlGhamdi, 2020). وتُعتبر الجبيل أحد أكبر المراكز الصناعية في المملكة، حيث تحتضن العديد من الصناعات البتروكيمياوية ومحطات الطاقة، مما يزيد من تركيز الملوثات في الغلاف الجوي (Kumar, 2022; Mujabar & Rao, 2018). وبالرغم من الجهود المبذولة ضمن مبادرة "السعودية الخضراء" ورؤية ٢٠٣٠ للحد من الانبعاثات وتحسين الاستدامة البيئية، فإن الحاجة إلى نظم مراقبة متطورة لا تزال ملحة لمتابعة ديناميكيات التلوث واتخاذ تدابير استباقية للحد من آثاره (Kahia et al., 2021).

شهدت السنوات الأخيرة تطوراً كبيراً في تطبيقات الذكاء الاصطناعي والاستشعار عن بُعد لرصد وتحليل جودة الهواء، خاصة في المناطق الصناعية ذات التحديات البيئية المعقدة مثل الدمام والجبيل. يوفر الاستشعار عن بُعد أدوات قوية لرصد التغيرات الزمنية والمكانية لجودة الهواء، حيث يُمكن للأقمار الصناعية، مثل Sentinel-5 المزود بمستشعر TROPOMI، قياس مستويات الملوثات الجوية بدقة عالية وعلى نطاق واسع (Loyola et al., 2020; Virghileanu et al., 2018). ومن خلال تحليل صور الأقمار الصناعية، يمكن تحديد مناطق التركيز العالي للملوثات وتتبع أنماط انتشارها بمرور الوقت، مما يساهم في تحسين استراتيجيات إدارة جودة الهواء (Mukundan et al., 2022).

علاوة على ذلك، أصبحت تقنيات الذكاء الاصطناعي، وخاصة خوارزميات التعلم الآلي مثل الغابات العشوائية وXGBoost، من الأدوات الأساسية في تحليل البيانات البيئية والتنبؤ بمستويات التلوث المستقبلية (Chen & Guestrin, 2016; Reddy & Parvathy, 2022; Victoriano et al., 2022a). وقد أظهرت العديد من الدراسات الحديثة أن النماذج القائمة على الذكاء الاصطناعي توفر دقة تنبؤية عالية مقارنةً بالأساليب التقليدية، حيث يمكنها معالجة كميات كبيرة من البيانات وتحديد الأنماط المعقدة التي تربط بين العوامل المناخية والملوثات الجوية (Makky & Eghrari; Bayhan et al., 2025; Chen et al., 2025; Sujatha et al., 2025). كما تُسهم هذه التقنيات في تطوير نظم إنذار مبكر تساعد في تقليل المخاطر الصحية المرتبطة بتلوث الهواء وتعزيز الاستدامة البيئية،

خاصة مع التوجهات الحديثة نحو استخدام الذكاء الاصطناعي التفسيري (Explainable AI) لتوضيح المتغيرات الأكثر تأثيراً في جودة الهواء (Ghimire et al., 2025). بناءً عليه، تبرز مساهمة هذه الدراسة في توظيف أحدث تقنيات الذكاء الاصطناعي (XGBoost) وربطها ببيانات الاستشعار عن بُعد لرصد وتحليل جودة الهواء في بيئة صناعية معقدة، بما يعزز من القيمة العلمية والعملية للنتائج في سياق الأدبيات الحديثة.

تتبع أهمية هذه الدراسة من الحاجة الملحة لفهم ديناميكيات تلوث الهواء في المدن الصناعية الكبرى بالمملكة العربية السعودية، مثل الدمام والجبيل، حيث يشكل تلوث الهواء تحدياً بيئياً وصحياً ذا آثار مباشرة على صحة الإنسان وجودة الحياة. تؤدي الانبعاثات الناتجة عن الأنشطة الصناعية الكثيفة، وحركة المرور المتزايدة، والتوسع الحضري السريع إلى ارتفاع مستويات الملوثات الهوائية، ما يفاقم من معدلات الإصابة بالأمراض التنفسية والقلبية ويهدد استدامة الأنظمة البيئية في المنطقة. ورغم الجهود الوطنية المبذولة في إطار رؤية المملكة ٢٠٣٠ ومبادرة "السعودية الخضراء" للحد من الانبعاثات وتحسين جودة الهواء، إلا أن هناك فجوة معرفية واضحة تتعلق برصد وتحليل التغيرات الزمنية والمكانية للملوثات، خاصة في ظل محدودية بيانات المحطات الأرضية التقليدية.

تكتسب الدراسة أهميتها العلمية والتطبيقية من اعتمادها على تقنيات الاستشعار عن بُعد ونماذج الذكاء الاصطناعي المتقدمة، ما يتيح رصدًا دقيقًا وشاملاً لجودة الهواء وتحليل العوامل المؤثرة عليها على مدى زمني طويل. وتوفر نتائج الدراسة قاعدة معرفية كمية يمكن الاستناد إليها في تطوير سياسات بيئية أكثر فعالية، وتوجيه استراتيجيات إدارة التلوث، وتعزيز برامج الإنذار المبكر، بما يدعم تحقيق أهداف التنمية المستدامة في المملكة. كما تسد هذه الدراسة فجوة بحثية مهمة في الأدبيات المحلية والإقليمية حول استخدام الذكاء الاصطناعي والاستشعار عن بُعد في مراقبة جودة الهواء في البيئات الصناعية، وتقدم نموذجاً يمكن تعميمه على مناطق حضرية وصناعية مماثلة في المنطقة العربية.

تهدف هذه الدراسة إلى تقييم وتتبع جودة الهواء في مدينتي الدمام والجبيل باستخدام تقنيات الاستشعار عن بُعد ونماذج الذكاء الاصطناعي، من خلال تحليل البيانات الزمنية لرصد التغيرات في تركيزات الملوثات، فضلاً عن دراسة العوامل البيئية المؤثرة على مستويات التلوث. ومن المتوقع أن تُسهم نتائج هذه الدراسة في تحسين فهم ديناميكيات تلوث الهواء في المناطق الصناعية، مما يدعم تطوير سياسات بيئية أكثر كفاءة لمكافحة التلوث وتحسين جودة الحياة في المملكة العربية السعودية.

٢- الدراسات السابقة

شهدت السنوات الأخيرة تطوراً ملحوظاً في توظيف تقنيات الاستشعار عن بُعد ونماذج الذكاء الاصطناعي في تحليل جودة الهواء والتنبؤ بمستويات التلوث، خاصة في المناطق الصناعية والحضرية ذات الكثافة السكانية العالية. تقليدياً، اعتمدت دراسات جودة الهواء على بيانات المحطات الأرضية، إلا أن محدودية التغطية المكانية والزمانية دفعت الباحثين إلى الاستفادة من صور الأقمار الصناعية مثل Sentinel-5P وMODIS، التي توفر بيانات عالية الدقة لرصد الملوثات الجوية على نطاق واسع (Loyola et al., 2018; Virghileanu et al., 2020; Mukundan et al., 2022).

أظهرت تطبيقات الاستشعار عن بُعد قدرة كبيرة على تتبع التغيرات الزمنية والمكانية لمستويات الملوثات الرئيسية، مثل $PM_{2.5}$ و SO_2 و NO_2 ، وساهمت في تحسين فهم ديناميكيات التلوث في البيئات الحضرية والصناعية (Stratoulis, 2024). ومع تطور الذكاء الاصطناعي، أصبح بالإمكان دمج البيانات الفضائية مع بيانات الأرصاد الجوية والبيانات الأرضية في نماذج تنبؤية متقدمة، مثل XGBoost و Random Forest والشبكات العصبونية العميقة (Deep Learning)، لتحسين دقة التنبؤ بمؤشر جودة الهواء (AQI) (Chen & Guestrin, 2016; Reddy & Parvathy, 2022b; Victoriano et al., 2022; Almaliki et al., 2023).

تشير دراسات حديثة إلى أن استخدام نماذج الذكاء الاصطناعي، خاصة الشبكات العصبية المتكررة (LSTM, GRU) والنماذج الهجينة (CNN-LSTM)، يحقق دقة تنبؤية تتجاوز ٩٥% في التنبؤ بقيم $PM_{2.5}$ وAQI، كما هو موضح في دراسات حديثة قارنت بين أداء GRU و CNN و XGBoost في تحليل بيانات ضخمة متعددة الأبعاد (Saikumar et al., 2025; Chen et al., 2025; Kaleeswari et al., 2022). وأكدت هذه الدراسات أن دمج البيانات الفضائية مع تقنيات الذكاء الاصطناعي يعزز من فعالية نظم الإنذار المبكر ويوفر أدوات كمية تدعم صانع القرار البيئي في رسم السياسات الاستباقية (Mohan et al., 2025; Al-Rawabdeh et al., 2025; Baruah & Mazumder, 2025).

علاوة على ذلك، أظهرت المراجعات المنهجية أن خوارزميات مثل Random Forest و XGBoost و LSTM هي الأكثر استخداماً وفعالية في مجال التنبؤ بتلوث الهواء، حيث تتيح معالجة العلاقات المعقدة بين العوامل المناخية والملوثات، وتوفر تنبؤات دقيقة على المستويين الزمني والمكاني (Zhang et al., 2025; Kaveh et al., 2025). كما أن التطورات الأخيرة في الذكاء الاصطناعي التفسيري (Explainable AI) تتيح تحديد المتغيرات الأكثر تأثيراً في جودة الهواء، ما يعزز من موثوقية النماذج ويساعد في توجيه السياسات البيئية بشكل أكثر فعالية (Ahmed et al., 2024).

بناءً على ذلك، تبرز هذه الدراسة في توظيفها منهجية متكاملة تجمع بين بيانات الاستشعار عن بُعد ونماذج الذكاء الاصطناعي المتقدمة لتحليل جودة الهواء في مدينتي الدمام والجبيل، ما يسد فجوة بحثية مهمة على المستوى المحلي والإقليمي، ويواكب أحدث الاتجاهات العالمية في هذا المجال.

٣- وصف منطقة الدراسة

تُعد مدينة الجبيل الواقعة على الساحل الشرقي للمملكة العربية السعودية واحدة من أهم المراكز الصناعية في العالم، حيث تحتضن الجبيل الصناعية، التي تُعد من أكبر المدن الصناعية المتكاملة عالمياً شكل (١). تشرف الهيئة الملكية للجبيل وينبع على تطويرها وإدارتها، وقد تم إنشاؤها ضمن خطة استراتيجية لتحفيز الاقتصاد الوطني وتعزيز قطاعي البتروكيماويات والصناعات الثقيلة.

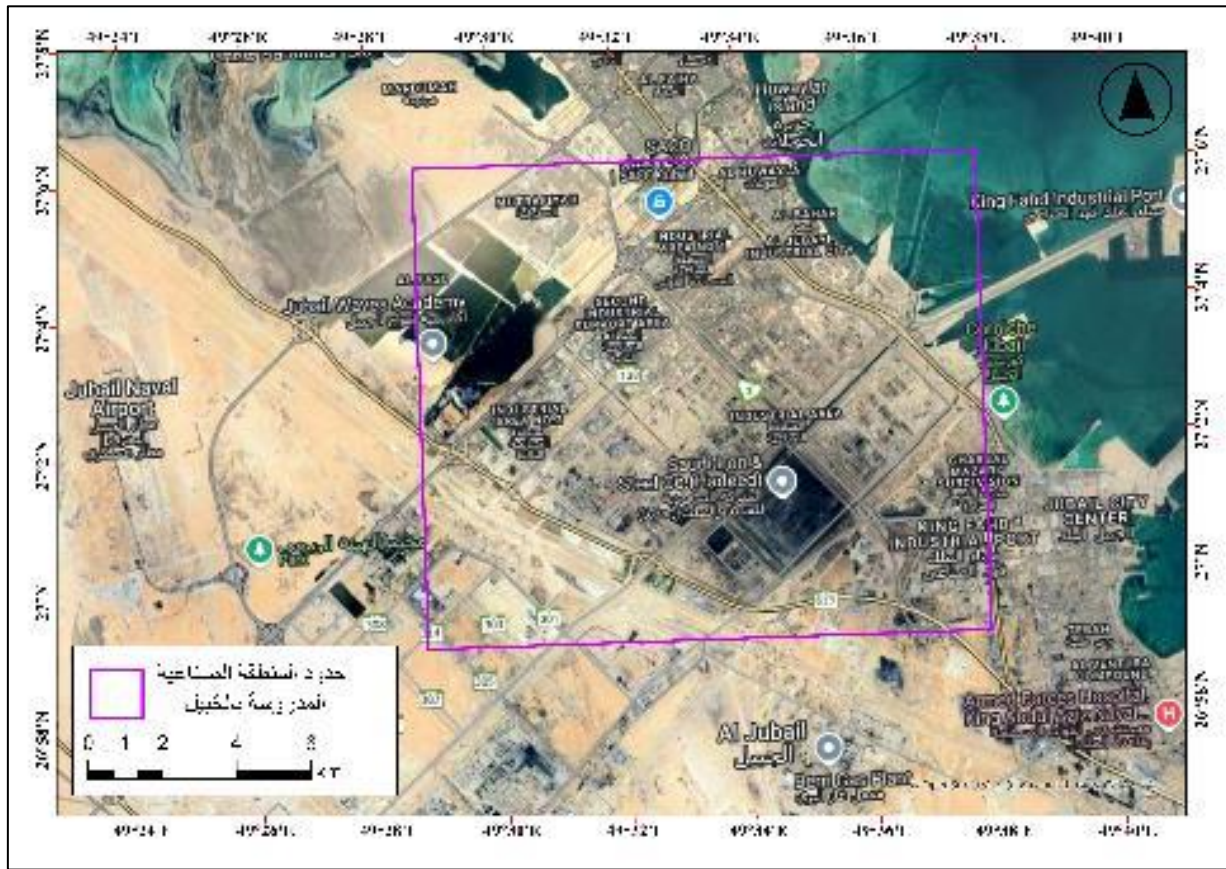
تضم الجبيل الصناعية العديد من المصانع العملاقة التي تنشط في مجالات مثل البتروكيماويات، التكرير، الحديد والصلب، وصناعة الأسمدة، مما جعلها مركزاً رئيسياً لجذب الاستثمارات الأجنبية والمحلية. من أبرز الشركات العاملة في المدينة: سابك، صدارة، التصنيع الوطنية، وشركة معادن، والتي تساهم في تحويل المدينة إلى محور صناعي عالمي.

إضافة إلى مدينة الدمام والتي تُعد عاصمة المنطقة الشرقية وأحد أهم المراكز الاقتصادية والصناعية في المملكة العربية السعودية شكل (٢). تقع على ساحل الخليج العربي وتتميز بموقع استراتيجي يجعلها مركزاً حيويًا للتجارة والصناعة والخدمات اللوجستية. تضم المدينة ميناء الملك عبد العزيز، الذي يُعد أكبر موانئ المملكة على الخليج العربي، مما يعزز مكانتها كحلقة وصل رئيسية في التجارة البحرية.

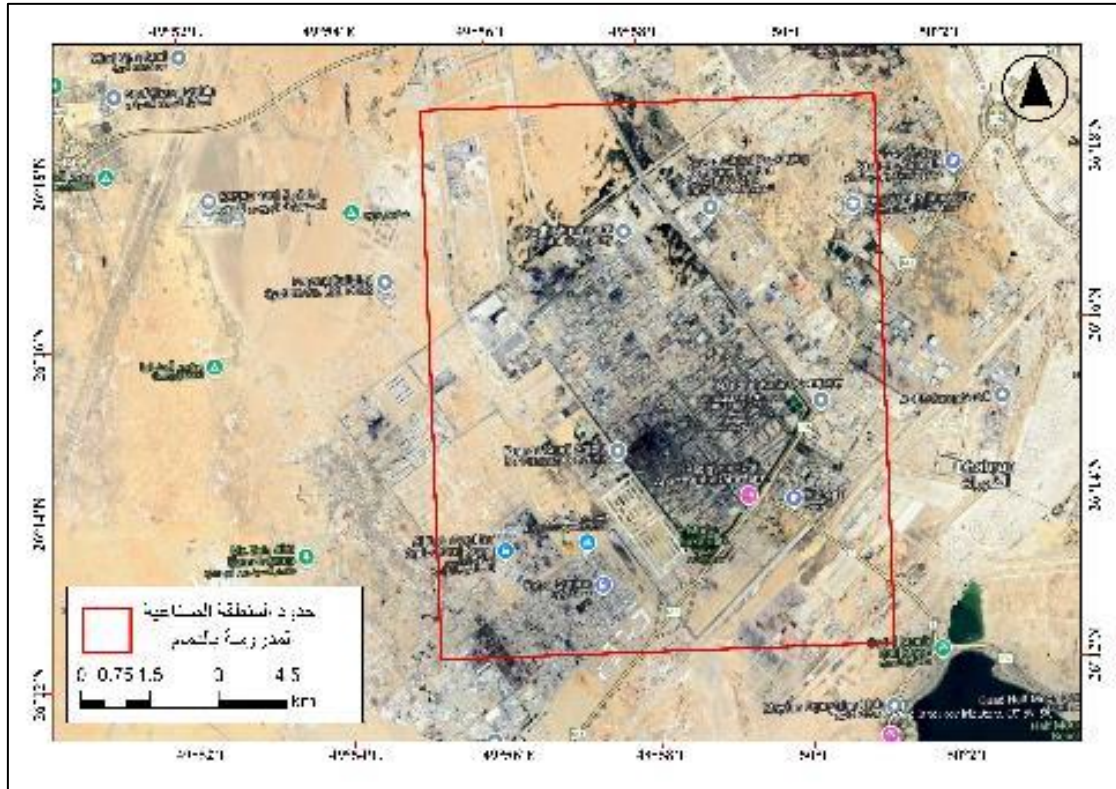
تُعتبر المنطقة الصناعية في الدمام من أكبر المناطق الصناعية في المملكة، حيث تضم المدينة الصناعية الأولى والثانية والثالثة، والتي تشرف عليها هيئة المدن الصناعية ومناطق التقنية (مدن). تستوعب هذه المناطق مجموعة متنوعة من الصناعات، تشمل:

- البتروكيماويات
- الصناعات التحويلية
- صناعة المعادن والحديد
- الأدوية
- الصناعات الغذائية

كما تستضيف المدينة عددًا من المصانع الكبرى مثل شركة أرامكو السعودية، والشركة السعودية للصناعات الأساسية (سابك)، وشركات الأسمنت والكيماويات، مما يجعلها أحد المحركات الرئيسية للاقتصاد الوطني. إلى جانب دورها الصناعي، تتمتع الدمام بتطور عمراني وبنية تحتية متقدمة، وتوفر بيئة متكاملة للمستثمرين والعاملين، مما يجعلها نقطة جذب رئيسية للمشاريع الاقتصادية الكبرى، وتساهم بشكل كبير في تحقيق رؤية المملكة ٢٠٣٠ من خلال تعزيز القطاع الصناعي وتنويع مصادر الدخل.



شكل ١. حدود المنطقة الصناعية المدروسة بالجبيل بالمملكة العربية السعودية.



شكل ٢. حدود المنطقة الصناعية المدروسة بالدمام بالمملكة العربية السعودية.

٤- المنهجية Methodology

٤-١ البيانات المستخدمة

تم جمع بيانات جودة الهواء باستخدام منصة Google Earth Engine (GEE)، التي تُعد من أقوى منصات الحوسبة السحابية لمعالجة وتحليل البيانات الجغرافية المكانية وبيانات الاستشعار عن بُعد (Haque et al., 2022). تتيج GEE الوصول إلى كميات هائلة من بيانات الأقمار الصناعية، مع إمكانية تنفيذ عمليات تحليلية معقدة باستخدام لغة JavaScript داخل محرر كود Earth Engine Code Editor. كما تتيج للمستخدمين الوصول إلى أدوات وخوارزميات متقدمة لتحليل بيانات الاستشعار عن بُعد بكفاءة عالية (Ghasempour et al., 2021).

في هذه الدراسة، تم الاعتماد على بيانات القمر الصناعي Sentinel-5P التابع لبرنامج Copernicus لمراقبة الغلاف الجوي، دون دمج بيانات من محطات رصد أرضية لمعايرة النتائج أو التحقق من دقتها. ويرجع ذلك إلى محدودية توفر بيانات محطات الرصد الأرضية في المنطقة المدروسة، حيث لا توجد شبكة رصد منتظمة تغطي جميع الملوثات الرئيسية بشكل متزامن وزمني كافٍ خلال فترة الدراسة. إضافة إلى ذلك، فإن المعايرة الدقيقة بين

بيانات الأقمار الصناعية وبيانات المحطات الأرضية تتطلب توافر بيانات متزامنة مكانياً وزمنياً لنفس الملوثات، وهو ما لم يكن متحققاً في المنطقة خلال فترة الدراسة. كما أن الفروق المنهجية بين ما تقيسه الأقمار الصناعية (العمود الكلي للملوثات) وما تسجله المحطات الأرضية (التركيز السطحي) تتطلب نماذج تحويل معقدة قد تقتقر للدقة في غياب بيانات أرضية كافية. وبناءً عليه، تم الاعتماد على منهجية الاستشعار عن بُعد المدعومة بنماذج الذكاء الاصطناعي، مع الإشارة الصريحة في الدراسة إلى أن النتائج تعكس الاتجاهات العامة لجودة الهواء، وتوفر أساساً علمياً يمكن البناء عليه مستقبلاً عند توفر بيانات أرضية أكثر شمولاً. ويوصى في الأبحاث القادمة بدمج بيانات المحطات الأرضية حال توفرها لتعزيز دقة النماذج والتحقق من النتائج بشكل أكثر تفصيلاً.

يوفر مستشعر TROPOMI بيانات عالية الدقة لمجموعة من الملوثات، بما في ذلك ثاني أكسيد النيتروجين NO_2 ، وثاني أكسيد الكبريت SO_2 ، وأول أكسيد الكربون CO ، والجسيمات الدقيقة $PM_{2.5}$ ، والأوزون O_3 (Loyola et al., 2018). تم استخراج البيانات اليومية للفترة الزمنية الممتدة بين يناير ٢٠١٩ إلى يونيو ٢٠٢٤، مع التركيز على مدينتي الدمام والجبيل باعتبارهما من المراكز الصناعية الرئيسية في المملكة العربية السعودية. وقد بلغت الدقة المكانية ١ كم.

علاوة على ذلك، تم استخدام بيانات من القمر الصناعي (MODIS (Moderate Resolution Imaging Spectroradiometer) لقياس معامل الانعكاسية الأرضية ودرجات حرارة سطح الأرض LST ، مما يساعد في تفسير العوامل البيئية المؤثرة على تباين جودة الهواء (Ghasempour et al., 2021). تمت معالجة البيانات وتحليلها باستخدام تقنيات تحليل السلاسل الزمنية والتصنيف المكاني، وذلك لتحديد الاتجاهات الزمنية والمكانية للملوثات الجوية وتقييم تأثير العوامل المناخية والمناطق الصناعية على جودة الهواء.

بشكل عام، يساهم استخدام GEE في تسريع عمليات تحليل البيانات الضخمة بفضل إمكانيات الحوسبة السحابية، مما يُمكن الباحثين من تحليل البيانات البيئية بدقة وكفاءة، ويوفر إطاراً متكاملًا للتنبؤ بمستويات تلوث الهواء اعتماداً على بيانات الاستشعار عن بُعد والذكاء الاصطناعي.

٢-٤ مؤشر جودة الهواء (AQI) Air Quality Index

مؤشر جودة الهواء AQI هو طريقة معيارية تُستخدم لنقل مستوى تلوث الهواء. يتم حسابه بناءً على تركيز الملوثات الرئيسية في الهواء، بما في ذلك $PM_{2.5}$ ، PM_{10} ، CO ، NO_2 ، SO_2 ، و O_3 . تعتمد الدول المختلفة أنظمة AQI خاصة بها، ولكن الحساب العام يتبع منهجية موحدة.

يتم حساب مؤشر AQI لكل من الملوثات الرئيسية التالية:

- الجسيمات الدقيقة $PM_{2.5}$
- أول أكسيد الكربون CO
- ثاني أكسيد النيتروجين NO_2
- ثاني أكسيد الكبريت SO_2
- الأوزون O_3

كل ملوث له قيم حدودية تُعرف بـ نقاط التوقف Breakpoints، وهي نطاقات تركيز مرتبطة بفئات AQI. تُوفر معظم الدول، بما في ذلك وكالة حماية البيئة الأمريكية US EPA ومنظمة الصحة العالمية (WHO)، جداول تحدد حدود AQI لكل ملوث. يوضح جدول (١) تصنيف جودة الهواء AQI بناءً على تركيز الملوثات الرئيسية طبقاً لنقاط التوقف Breakpoints.

نقاط التوقف Breakpoints تعتمد على نظام وكالة حماية البيئة الأمريكية (US EPA). من خلال نقاط التوقف يمكن حساب مؤشر فرعي لهذه التركيزات.

جدول ١. تصنيف جودة الهواء AQI بناءً على تركيز الملوثات الرئيسية.

نطاق AQI	$PM_{2.5}$ ($\mu g/m^3$)	CO (ppm)	NO_2 (ppb)	SO_2 (ppb)	O_3 (ppm)	فئة جودة الهواء
0-50	0.0 - 12.0	0.0 - 4.4	0 - 53	0 - 35	0 - 54	جيد (Good)
51-100	12.1 - 35.4	4.5 - 9.4	54 - 100	36 - 75	55 - 70	مُرَضِي (Satisfactory)
101-150	35.5 - 55.4	9.5 - 12.4	101 - 360	76 - 185	71 - 85	معتدل (Moderate)
151-200	55.5 - 150.4	12.5 - 15.4	361 - 649	186 - 304	86 - 105	سيء (Poor)
201-300	150.5 - 250.4	15.5 - 30.4	650 - 1249	305 - 604	106 - 200	سيء للغاية (Very Poor)
301-500	250.5 - 500.4	30.5 - 50.4	1250 - 2049	605 - 1004	>200	Severe (خطير)

المصدر: (<https://www.airnow.gov/aqi/aqi-basics>)

يتم حساب AQI لأي ملوث باستخدام المعادلة التالية:

$$AQI = \left(\frac{I_H - I_L}{C_H - C_L} \right) \times (C - C_L) + I_L$$

AQI = Air Quality Index for pollutant

C = Measured concentration of the pollutant

C_L = Lower bound of the concentration range in the AQI category

C_H = Upper bound of the concentration range in the AQI category

I_L = Lower bound of the AQI value corresponding to C_L

I_H = Upper bound of the AQI value corresponding to C_H

يتم تحديد AQI الإجمالي لموقع معين من خلال أعلى قيمة AQI بين جميع الملوثات كما يلي:

$$AQI = \max (AQI_{PM2.5}, AQI_{CO}, AQI_{NO2}, AQI_{SO2}, AQI_{O3})$$

إذا كان أحد الملوثات لديه AQI أعلى من الآخرين، فإنه يحدد مستوى جودة الهواء العام.

٣-٤ نموذج XGBoost

يُعتبر XGBoost اختصار لـ eXtreme Gradient Boosting أحد أكثر خوارزميات التعلم الآلي فاعلية وقوة في معالجة الانحدار والتصنيف. يعتمد XGBoost على إطار عمل Gradient Boosting، حيث يتم بناء النموذج بشكل تسلسلي من خلال تحسين الأخطاء Residuals للنماذج السابقة باستخدام تقنية تعزيز التدرج Gradient Descent. يتميز XGBoost بقدرته على التعامل مع البيانات الكبيرة والمعقدة، فضلاً عن دعمه للتوازي في الحسابات، مما يجعله أسرع وأكثر كفاءة مقارنة بالخوارزميات التقليدية. (Chen & Guestrin, 2016). وقد تم استخدام هذا النموذج في التنبؤ المستقبلي لقيم جودة الهواء للفترة (٢٠٢٥-٢٠٣٠).

يتميز XGBoost بعدة خصائص تقنية تجعله متفوقاً على العديد من الخوارزميات التقليدية:

- **الكفاءة والسرعة:** يدعم XGBoost التنفيذ المتوازي (Parallelization) أثناء بناء الأشجار، ما يقلل من زمن التدريب بشكل كبير، ويجعله مناسباً للتعامل مع مجموعات البيانات الضخمة والمعقدة.
- **المرونة:** يمكن تخصيص دوال الخسارة والمعاملات الفائقة (Hyperparameters) بما يتناسب مع طبيعة المشكلة، سواء كانت تصنيفية أو انحدارية. كما يدعم النموذج التعامل مع القيم المفقودة تلقائياً، ما يجعله عملياً في البيئات الحقيقية.
- **مقاومة الإفراط في الملاءمة:** يوفر XGBoost تقنيات تنظيم (Regularization) متقدمة مثل $L1$ و $L2$ ، والتي تساعد في تقليل خطر الإفراط في ملاءمة النموذج للبيانات التدريبية، وتحسين قدرة النموذج على التعميم عند تطبيقه على بيانات جديدة.
- **الدعم البرمجي والتكامل:** يتوفر XGBoost كمكتبة مفتوحة المصدر تدعم لغات برمجة متعددة مثل Python و R و Java، ويتكامل بسهولة مع مكتبات علم البيانات مثل scikit-learn. ولتحسين أداء النموذج، يتم إجراء ضبط المعلمات الفائقة باستخدام Randomized Search (CV)، وتحسين المعلمات مثل عدد المقدرين ومعدل التعلم وعمق الشجرة ونسب أخذ العينات من الميزات. يتم اختيار النموذج

الأفضل أداءً بناءً على أدنى خطأ متوسط مربع. بمجرد التدريب، يتم تقييم قدرة النموذج التنبؤية باستخدام مقاييس التقييم القياسية، بما في ذلك خطأ متوسط مربع MSE ومعامل التحديد R^2 ، والتي توفر رؤى حول دقة التنبؤ. تم استخدام هذا النموذج في التنبؤ المستقبلي بمؤشر جودة الهواء للفترة (٢٠٢٥-٢٠٣٠). وذلك باستخدام نهج قائم على التأخير، مع دمج القيم التاريخية كخصائص تنبؤية. يتم تطبيق تعديل موسمي من خلال الاستفادة من متوسط الاتجاهات الشهرية والانحراف المعياري لتحسين التوقعات.

وقد تم ضبط المعلمات الفائقة لـ XGBoost كما يلي:

```
param_grid = {
    'n_estimators': [100, 200, 300, 400, 500, 600, 700],
    'max_depth': [1, 3, 5, 4],
    'learning_rate': [0.01, 0.05, 0.1],
    'subsample': [0.1, 0.5, 1.0, 0.8, 0.9],
    'colsample_bytree': [0.1, 0.9, 1.0, 0.1, 0.01]
```

وتم استخدام RandomizedSearchCV كما يلي:

```
random_search = RandomizedSearchCV(
    estimator=xgb_model,
    param_distributions=param_grid,
    n_iter=10, # Number of parameter combinations to try
    cv=5,
    scoring='neg_mean_squared_error',
    verbose=2,
    random_state=42,
    n_jobs=-1 # Use all available CPU cores
)
```

❖ تقييم أداء النموذج باستخدام ثلاثة مقاييس رئيسية:

- متوسط الخطأ المطلق (MAE): Mean Absolute Error (MAE): يقيس متوسط الفروق المطلقة بين القيم الحقيقية والقيم المتوقعة، مما يساعد في تقييم دقة النموذج.

- متوسط خطأ الجذر التربيعي (Root Mean Squared Error (RMSE): يحدد مدى اختلاف القيم المتوقعة عن القيم الفعلية، مما يعكس حجم الخطأ.
- معامل التحديد (Coefficient of Determination (R^2): يوضح نسبة التباين في البيانات التي يمكن للنموذج تفسيرها، حيث تشير القيم الأقرب إلى ١ إلى أداء تنبؤي أفضل.
- تحليل المخلفات (Residuals) للتحقق من جودة التنبؤات.

٤-٤ انحدار ثيل-سين *Theil-Sen Estimator*

يُعتبر انحدار ثيل-سين (Theil-Sen Estimator) أسلوبًا إحصائيًا غير معلمي يُستخدم لتقدير خط الانحدار في تحليل الانحدار الخطي البسيط، وذلك من خلال حساب ميل الخط كوسيط لميل جميع الخطوط المارة عبر أزواج من النقاط في العينة. يُعرف هذا المقدّر بقدرته على مقاومة القيم الشاذة (outliers)، مما يجعله أكثر دقة مقارنةً بأساليب الانحدار التقليدية، خاصةً عند التعامل مع بيانات منحرفة أو غير متجانسة (Sen, 1968). كما يُعتبر من أكثر التقنيات غير المعلمية شيوعًا لتقدير الاتجاهات الخطية في البيانات البيئية والاقتصادية. وتم استخدامه لتحديد اتجاه الملوثات الجوية بالجبل والدمام وكذلك مؤشر جودة الهواء.

٤-٥ البرامج المستخدمة

تم الاعتماد على برنامج Visual Studio Code وذلك لكتابة الاكواد وتحليل البيانات واستخراج الاشكال البيانية وتطبيق نموذج XGBoost بالاعتماد على لغة البايثون Python الإصدار ٣,٩ ومكتباتها مثل:

- numpy & pandas لمعالجة البيانات.
 - scikit-learn لضبط النماذج والتحقق المتقاطع.
 - xgboost لبناء نموذج XGBoost.
 - matplotlib و seaborn للرسم البياني وتحليل المخلفات.
- بالاضافة إلى RandomizedSearchCV لضبط معاملات نموذج XGBoost واختيار النموذج الأمثل بناءً على أقل خطأ تربيعي.

تم الاعتماد على برنامج Arc GIS Pro v.3 وذلك لإنتاج وإخراج الخرائط الخاصة بتوزيع الملوثات.

٥- نتائج الدراسة

٥-١ التحليل الإحصائي لتركيزات الملوثات الجوية بالجبيل والدمام

يكشف التقييم المقارن لمستويات تلوث الهواء في الجبيل والدمام عن اختلافات واضحة في تركيزات الملوثات الجوية الرئيسية، بما في ذلك الجسيمات العالقة ($PM_{2.5}$)، وأول أكسيد الكربون (CO)، وثاني أكسيد النيتروجين (NO_2)، وثاني أكسيد الكبريت (SO_2)، والأوزون (O_3). تسلط هذه الاختلافات الضوء على التأثير المحتمل للأنشطة الصناعية، وانبعاثات المركبات، والعوامل الجوية على جودة الهواء في كلتا المدينتين.

يُظهر $PM_{2.5}$ ، وهو مؤشر رئيسي لتلوث الجسيمات الدقيقة، تركيزات أعلى في الجبيل، بمتوسط ٤٥,٢٨ ميكروجرام/م^٣، مقارنة بـ ٤٢,٠٤ ميكروجرام/م^٣ في الدمام. تتجاوز مستويات $PM_{2.5}$ القصوى المسجلة في الجبيل (٤٥٨,٦٧ ميكروجرام/م^٣) تلك الموجودة في الدمام (٣٧٤,٩١ ميكروجرام/م^٣)، مما يشير إلى حدوث حالات تلوث شديدة، ربما مرتبطة بالانبعاثات الصناعية وحركة المركبات والظروف الجوية مثل الركود الجوي. يشير الانحراف المعياري الأعلى في الجبيل (٣٠,٦٤ ميكروجرام/م^٣) إلى تباين أكبر في مستويات التلوث، والذي قد يتأثر بتقلبات الأنشطة الصناعية وأنماط الرياح وانعكاسات درجات الحرارة التي تحبس الملوثات بالقرب من السطح (جدول ٢).

يشير تحليل تركيزات أول أكسيد الكربون إلى أن مستويات التلوث تظل مستقرة نسبياً في كلتا المدينتين. إن متوسط تركيز أول أكسيد الكربون في الدمام (٣,٥٢ جزء في المليون) أعلى قليلاً من تركيزه في الجبيل (٣,٣٧ جزء في المليون)، لكن الفارق ضئيل. وبالمثل، فإن مستويات أول أكسيد الكربون القصوى في الدمام (٤,٧٤ جزء في المليون) والجبيل (٤,٥٥ جزء في المليون) متقاربة، مما يشير إلى أن انبعاثات أول أكسيد الكربون تنشأ في المقام الأول من احتراق المركبات وحرق الوقود الصناعي في كلا الموقعين. تؤكد الانحرافات المعيارية المنخفضة (٠,٣٠ جزء في المليون في الدمام و٠,٣٣ جزء في المليون في الجبيل) أن تركيزات أول أكسيد الكربون لا تظهر تقلبات كبيرة بمرور الوقت، مما يعزز فكرة أن الانبعاثات ثابتة نسبياً ولا تتأثر بشكل كبير بالأنشطة الصناعية المتفرقة أو التقلبات الجوية.

جدول ٢. إحصائيات مقارنة لتلوث الهواء في الجبيل والدمام.

الدمام	الجبيل	المعاملات الإحصائية	الملوثات
42.04	45.28	Average ($\mu g/m^3$)	$PM_{2.5}$
10.82	12.28	Minimum	
374.91	458.67	Maximum	
24.79	30.64	Standard Deviation	
3.52	3.37	Average (ppm)	
2.51	2.14	Minimum	

CO	Maximum	4.55	4.74
	Standard Deviation	0.33	0.30
	Average (ppb)	13.77	12.41
NO ₂	Minimum	0.96	1.18
	Maximum	103.92	59.90
	Standard Deviation	9.59	7.46
SO ₂	Average (ppb)	32.38	26.48
	Minimum	-96.70	-85.80
	Maximum	544.18	351.67
	Standard Deviation	43.39	39.59
O ₃	Average (ppm)	12.49	12.42
	Minimum	9.43	9.46
	Maximum	15.40	15.29
	Standard Deviation	0.68	0.68

يُظهر أكسيد النيتروجين، وهو أحد الأسباب الرئيسية للضباب الدخاني الحضري والملوثات الثانوية مثل الأوزون، اختلافات ملحوظة بين المدينتين. يتجاوز متوسط تركيز أكسيد النيتروجين في الجبيل (١٣,٧٧ جزء في المليار) تركيزه في الدمام (١٢,٤١ جزء في المليار)، والذي قد يُعزى إلى الانبعاثات الصناعية الأعلى وحركة المرور الكثيفة في المنطقة. يُلاحظ فرق ملحوظ في مستويات أكسيد النيتروجين القصوى، حيث تصل الجبيل إلى ١٠٣,٩٢ ميكروجرام/م^٣ - وهو ما يقرب من ضعف ٥٩,٩٠ ميكروجرام/م^٣ المسجلة في الدمام. يشير هذا إلى أن أحداث التلوث المحلية، المرتبطة ربما بعمليات الاحتراق الصناعي أو المناطق ذات حركة المرور الكثيفة، تساهم في مستويات أكسيد النيتروجين الشديدة في الجبيل. يشير الانحراف المعياري الأعلى في الجبيل (٩,٥٩ جزء في المليار) أيضًا إلى تقلبات أكثر وضوحًا في انبعاثات أكسيد النيتروجين، والتي يمكن أن تتأثر بالاختلافات في كفاءة احتراق الوقود، وأنماط التشتت الجوي، والتفاعلات الكيميائية الناجمة عن درجة الحرارة.

يُظهر ثاني أكسيد الكبريت، وهو ملوث ينبعث في المقام الأول من العمليات الصناعية مثل تكرير النفط وتوليد الطاقة وصهر المعادن، تباينًا كبيرًا بين المدينتين. إن متوسط تركيز ثاني أكسيد الكبريت في الجبيل (٣٢,٣٨ جزء في المليار) أعلى منه في الدمام (٢٦,٤٨ جزء في المليار)، مما يعكس تأثير الأنشطة الصناعية واسعة النطاق في الجبيل. والأمر الأكثر أهمية هو أن الحد الأقصى لمستوى ثاني أكسيد الكبريت في الجبيل (٥٤٤,١٨ جزء في المليار) أعلى بكثير من الدمام (٣٥١,٦٧ جزء في المليار)، مما يشير إلى أن بعض المناطق الصناعية في الجبيل قد تشهد ارتفاعات متقطعة في انبعاثات ثاني أكسيد الكبريت، ربما بسبب الانبعاثات غير المنظمة أو عدم كفاءة العمليات. من المحتمل أن يكون وجود قيم أدنى سلبية لثاني أكسيد الكبريت (-٩٦,٧٠ جزء في المليار في الجبيل و-٨٥,٨٠ جزء في المليار في الدمام) ناتجًا عن أخطاء معايرة الأجهزة أو الظروف

الجوية التي تؤثر على انتشار الملوثات، وليس عن تركيزات ثاني أكسيد الكبريت السلبية الفعلية. تشير الانحرافات المعيارية المرتفعة نسبياً (٤٣,٣٩ جزء في المليار في الجبيل و ٣٩,٥٩ جزء في المليار في الدمام) إلى تقلبات زمنية كبيرة في مستويات ثاني أكسيد الكبريت، ربما بسبب دورات الإنتاج الصناعي والاختلافات في الظروف الجوية. فيما يتعلق بظهور قيم سالبة لتركيزات ثاني أكسيد الكبريت (SO_2) في نتائج الدراسة، نود التوضيح أن هذه القيم لا تعكس وجود تركيزات سالبة فعلية لهذا الغاز في البيئة، إذ أن ذلك غير ممكن علمياً. إنما ترجع هذه القيم السالبة إلى حدود ودقة تقنيات القياس والمعالجة، سواء في بيانات الأقمار الصناعية أو في النماذج الإحصائية المستخدمة. وقد أشارت العديد من الدراسات مثل (Khreis et al., 2022) إلى أن مثل هذه القيم قد تنتج عن عوامل مثل انحراف معايرة المستشعرات، أو التداخل مع غازات أخرى، أو تقنيات تصحيح الضوضاء في البيانات، خاصة عندما تكون التركيزات الحقيقية منخفضة جداً أو قريبة من حد الكشف للأجهزة. بناءً عليه، تم التعامل مع هذه القيم بحذر ضمن التحليل الإحصائي، مع الإشارة إلى أنها لا تؤثر على الاتجاهات العامة أو الاستنتاجات الرئيسية للدراسة، وسيتم مستقبلاً استبعادها أو تصحيحها في حال توفر بيانات أرضية معايرة أو نماذج تصحيح أكثر دقة.

وعلى النقيض من الملوثات الأخرى، تظل تركيزات O_3 متطابقة تقريباً في كلتا المدينتين، حيث سجلت الجبيل متوسطاً قدره ١٢,٤٩ جزء في المليون والدمام عند ١٢,٤٢ جزء في المليون. والقيم القصوى متقاربة أيضاً، عند ١٥,٤٠ جزء في المليون في الجبيل و ١٥,٢٩ جزء في المليون في الدمام، مما يشير إلى عمليات تكوين الأوزون الضوئية الكيميائية المتسقة. ونظراً لأن O_3 هو ملوث ثانوي يتكون من تفاعل NO_2 مع المركبات العضوية المتطايرة (VOCs) في وجود ضوء الشمس، فمن المتوقع أن تكون تركيزاته موحدة نسبياً عبر البيئات الحضرية ذات مصادر الانبعاثات المماثلة. ويدعم الانحراف المعياري المنخفض (٠,٦٨ جزء في المليون في كلتا المدينتين) الاستنتاج القائل بأن مستويات الأوزون تظل مستقرة بمرور الوقت، مع القليل من التأثير من الاختلافات في الانبعاثات قصيرة الأجل.

وتشير نتائج هذا التحليل إلى أن مدينة الجبيل تشهد عموماً مستويات أعلى من تلوث الهواء مقارنة بالدمام، وخاصة من حيث تركيزات $PM_{2.2}$ و NO_2 و SO_2 . ويشير وجود مستويات أعلى من الملوثات القصوى والتباين الأكبر في الجبيل إلى تأثير الأنشطة الصناعية الكبرى، والانبعاثات الثقيلة من المركبات، وارتفاعات التلوث العرضية. وفي حين تظل مستويات ثاني أكسيد الكربون والأكسجين مستقرة وقابلة للمقارنة في كلتا المدينتين، فإن التركيزات المرتفعة للجسيمات والملوثات القائمة على الكبريت تسلط الضوء على الحاجة إلى استراتيجيات إدارة جودة الهواء المستهدفة في الجبيل.

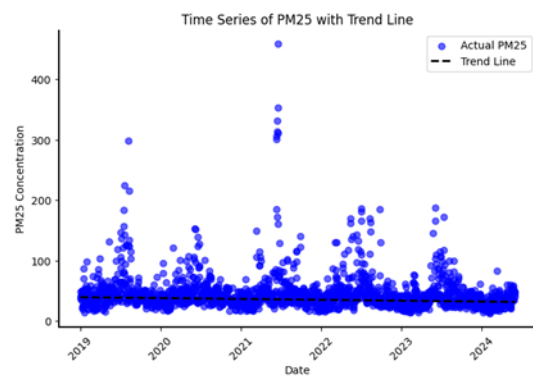
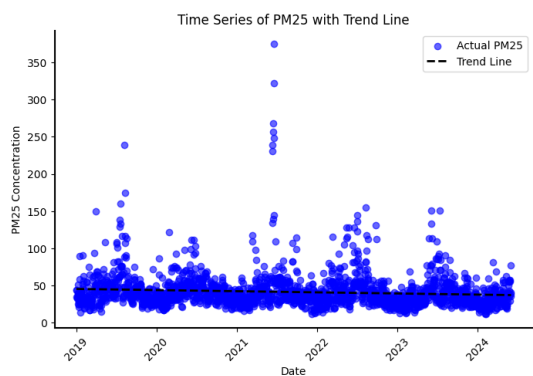
٥-٢ اتجاهات السلاسل الزمنية لتركيزات الملوثات الجوية بالجبيل والدمام

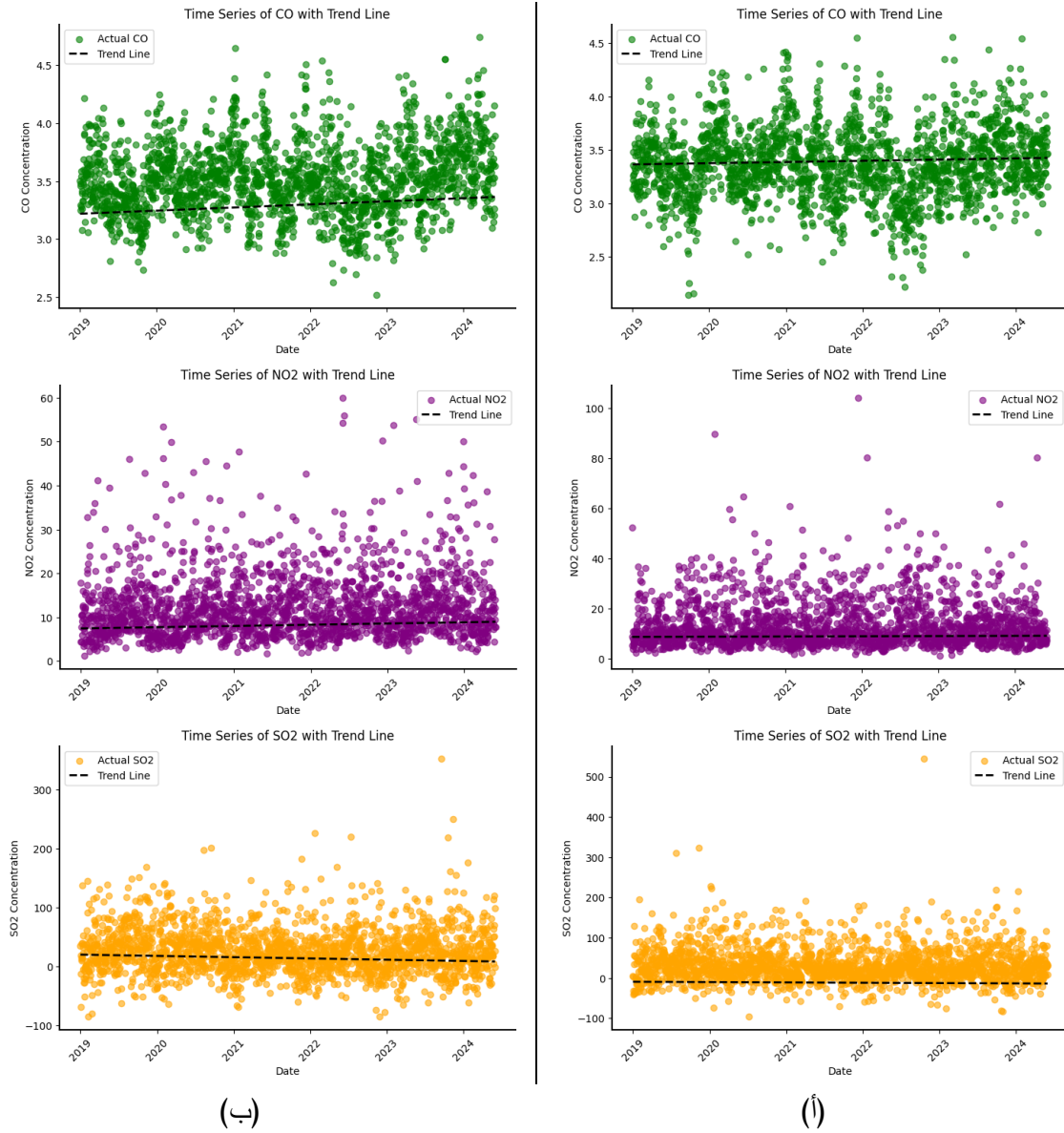
يوضح جدول (٣) وشكل (٣) الاتجاهات الزمنية للملوثات الجوية في مدينتي الجبيل والدمام باستخدام انحدار Theil-Sen ، وهو نموذج إحصائي قوي يستخدم لاكتشاف الاتجاهات في البيانات البيئية، نظرًا لقدرته على التعامل مع القيم المتطرفة. يشمل التحليل خمسة ملوثات رئيسية $PM_{2.5}$ ، CO ، NO_2 ، SO_2 ، O_3 ، وتم تقدير معامل الانحدار Slope، Intercept، والقيمة الاحتمالية P-value لكل ملوث في كل مدينة.

تشير نتائج التحليل إلى اتجاه تنازلي واضح في تركيز الجسيمات الدقيقة $PM_{2.5}$ في كل من الجبيل والدمام، حيث بلغ معامل الانحدار -٠,٠٠٣٨٨٥ و -٠,٠٠٤١٨٢ على التوالي. هذه القيم تدل على انخفاض تركيز $PM_{2.5}$ بمرور الوقت، وهو ما يعكس تأثير التدابير البيئية التي تهدف إلى تقليل مستويات التلوث الناجم عن الأنشطة الصناعية والانبعاثات المرورية. كما أن القيمة الاحتمالية ($P\text{-value} = 0.0001$) تشير إلى أن هذا الاتجاه ذو دلالة إحصائية عالية، مما يؤكد موثوقية النتائج.

جدول ٣. اتجاه الملوثات الجوية في الجبيل والدمام باستخدام انحدار Theil-Sen.

المتغيرات	الجبيل			الدمام		
	Slope	Intercept	P-value	Slope	Intercept	P-value
$PM_{2.5}$	-0.003885	2902.790723	0.0001	-0.004182	3127.61522	0.0001
CO	0.000031	-19.414836	0.0679	0.000073	-50.931597	0.0001
NO_2	0.000230	-160.806122	0.775	0.000767	-557.603333	0.0079
SO_2	-0.002208	1618.195269	0.026	-0.005877	4351.642689	0.0232
O_3	0.000191	-127.686726	0.0001	0.000164	-108.222035	0.0001





شكل ٣. اتجاهات السلاسل الزمنية لتركيزات الملوثات الجوية. (أ) الجبيل، (ب) الدمام.

فيما يخص أول أكسيد الكربون CO، لوحظ اتجاه تصاعدي طفيف في الجبيل، حيث بلغ معامل الانحدار 0.000031 ، بينما كان الاتجاه أكثر وضوحاً في الدمام بواقع 0.000073 . ومع ذلك، فإن القيمة الاحتمالية للجبيل ($P\text{-value} = 0.0679$) تشير إلى أن هذا الاتجاه ليس دالاً إحصائياً، على عكس الدمام حيث كانت القيمة الاحتمالية 0.0001 ، مما يعكس زيادة مؤكدة في تركيز CO هناك. قد يكون هذا الارتفاع ناتجاً عن زيادة الأنشطة الصناعية والانبعثات الصادرة عن وسائل النقل.

أما بالنسبة لـ ثاني أكسيد النيتروجين NO_2 ، فإن معامل الانحدار في الجبيل كان منخفضاً جداً (٠,٠٠٠٢٣٠)، كما أن القيمة الاحتمالية المرتفعة ($P\text{-value} = 0.775$) تشير إلى عدم وجود اتجاه واضح. في المقابل، أظهرت الدمام اتجاهًا تصاعديًا ملحوظًا مع معامل انحدار قدره ٠,٠٠٠٧٦٧ وقيمة احتمالية ($P\text{-value} = 0.0079$)، مما يعني وجود ارتفاع معنوي في تركيز NO_2 مع مرور الوقت. يعكس ذلك تأثير التوسع العمراني وزيادة الانبعاثات من المركبات ومحطات الطاقة.

من ناحية أخرى، أظهر ثاني أكسيد الكبريت SO_2 اتجاهًا تنازليًا في كل من الجبيل والدمام، حيث بلغت قيم معامل الانحدار -٠,٠٠٢٢٠٨ و -٠,٠٠٥٨٧٧ على التوالي، مما يشير إلى انخفاض في تركيز SO_2 بمرور الوقت. القيمة الاحتمالية ($P\text{-value} = 0.026$) في الجبيل و ٠,٠٢٣٢ في الدمام (تدل على أن هذا الانخفاض ذو دلالة إحصائية، مما يعكس نجاح التدابير البيئية في تقليل انبعاثات الكبريت الناتجة عن العمليات الصناعية وحرق الوقود الأحفوري).

أخيرًا، بالنسبة لـ الأوزون O_3 ، فقد لوحظ اتجاه تصاعدي طفيف في كل من الجبيل (٠,٠٠٠١٩١) والدمام (٠,٠٠٠١٦٤)، وكانت القيم الاحتمالية منخفضة جدًا ($P\text{-value} = 0.0001$) مما يدل على زيادة مؤكدة في تركيز الأوزون مع مرور الوقت. يعكس هذا الاتجاه التأثيرات الثانوية الناجمة عن التفاعلات الكيميائية بين الملوثات الأخرى مثل NO_2 والمركبات العضوية المتطايرة (VOCs) في وجود أشعة الشمس، مما يؤدي إلى تكوين الأوزون على مستوى سطح الأرض.

٣-٥ التوزيع المكاني لتركيزات الملوثات الجوية بالجبيل والدمام

يوضح شكل (٤) التوزيع المكاني لمتوسط الملوثات الجوية بمنطقة الدراسة للفترة (يناير ٢٠١٩ - يونيو ٢٠٢٤)، يمثل (أ-٤) الجسيمات العالقة $PM_{2.5}$ ، (ب-٤) الأوزون O_3 ، (ج-٤) أول أكسيد الكربون CO، (د-٤) ثاني أكسيد الكبريت SO_2 ، (هـ-٤) ثاني أكسيد النيتروجين NO_2 . تتراوح تركيزات $PM_{2.5}$ في المنطقة المدروسة بين ٣٧,١٤١ و ٥٢,٨١٨ ميكروغرام لكل متر مكعب. هذه القيم تشير إلى مستويات عالية من التلوث الجوي، حيث إن الحد الأقصى المسموح به من قبل منظمة الصحة العالمية هو ٢٥ ميكروغرام لكل متر مكعب على مدى ٢٤ ساعة. ويلاحظ أن منطقة الجبيل يزداد بها وجود تركيزات $PM_{2.5}$ عن منطقة الدمام. وكذلك في تركيزات الأوزون O_3 ، تزداد في الجبيل عن الدمام وتتراوح القيم ما بين ١٢٤,١ إلى ١٢٦,٢ جزء في المليون. بينما حدث العكس في تركيزات أول أكسيد الكربون CO، فيلاحظ أن الدمام يزداد بها تركيزات هذا الغاز عن الجبيل، وتتراوح قيمه ما بين ٣٠,٣ إلى ٣٥,٨ جزء في المليون. بينما ثاني أكسيد الكبريت SO_2 تتراوح قيمه ما بين ١٥١,٣ إلى

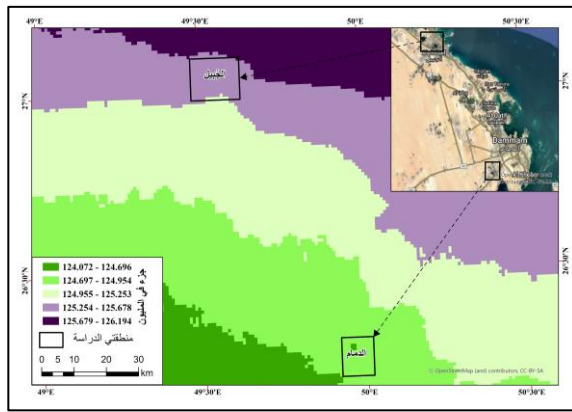
٦٠٠ جزء في المليار، ويتركز في الجبيل عن الدمام. أما ثاني أكسيد النيتروجين NO_2 فتزداد قيمه في كلتا المنطقتين وتتراوح القيم ما بين ٢١,٧ إلى ١٨٠,٣ جزء في المليار.

٥-٤ اتجاه مؤشر جودة الهواء وتصنيفه بالجبيل والدمام

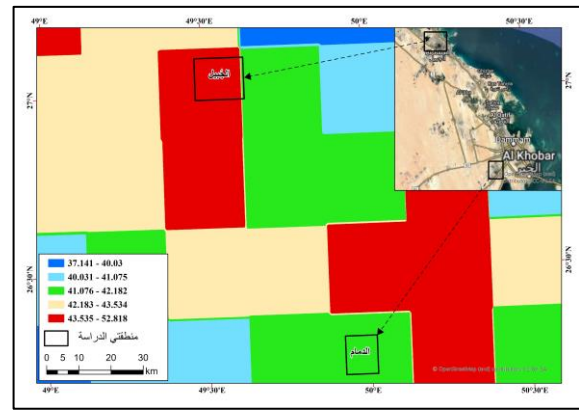
يعكس جدول (٤) الاتجاهات الزمنية لتصنيف جودة الهواء في الدمام والجبيل من عام ٢٠١٩ إلى عام ٢٠٢٤ وفقاً لمؤشرات AQI مؤشر جودة الهواء. ويشمل التصنيفات المختلفة: جيد، مُرضٍ، معتدل، سيئ، وسيئ جداً. يشير الجدول إلى أن الفئة "المُرضية" كانت الأكثر شيوعاً في كلا المدينتين خلال فترة الدراسة، حيث شهدت الجبيل زيادة في عدد الأيام المصنفة ضمن هذه الفئة من ١٠٤ يوماً في عام ٢٠١٩ إلى ١٥٣ يوماً في عام ٢٠٢٣، بينما سجلت الدمام ارتفاعاً من ١١٨ يوماً في عام ٢٠١٩ إلى ١٧٦ يوماً في عام ٢٠٢٣، مما يشير إلى تحسن نسبي في جودة الهواء.

بالنسبة لفئة "المعتدل"، فقد شهدت أيضاً انخفاضاً في كلتا المدينتين على مدار السنوات. ففي الجبيل، انخفض عدد الأيام المصنفة كمعتدلة من ٢٥٤ يوماً في ٢٠١٩ إلى ٢٠٩ يوماً في ٢٠٢٣، وفي الدمام انخفض العدد من ٢٤٤ يوماً إلى ١٨٦ يوماً خلال نفس الفترة. هذا الانخفاض قد يكون ناتجاً عن تحسن في التحكم في مصادر التلوث الصناعي أو زيادة وعي السكان بتأثيرات التلوث الهوائي.

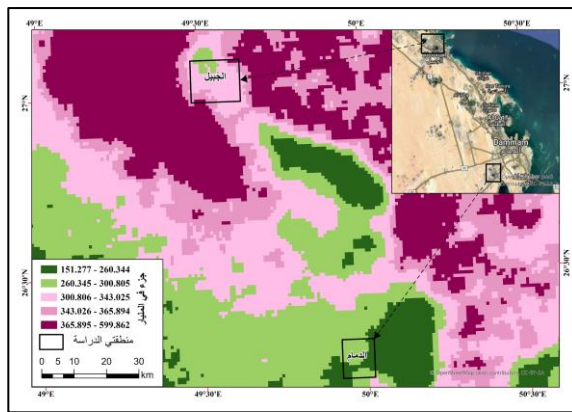
أما فيما يتعلق بالفئات الأكثر تدهوراً لجودة الهواء، مثل "السيئ" و"السيئ جداً"، فإن عدد الأيام المصنفة ضمن هذه الفئات ظل منخفضاً بشكل عام في كلا المدينتين. في الدمام، لم تسجل أي أيام مصنفة ضمن فئة "السيئ جداً" خلال الفترة المدروسة، بينما في الجبيل، كان هناك انخفاض ملحوظ حيث لم تسجل أي أيام ضمن هذه الفئة بعد عام ٢٠٢١، حيث سجلت ٦ أيام في عام ٢٠٢١ وانخفضت إلى ٠ في الأعوام اللاحقة. كما أن عدد الأيام التي تم تصنيفها ضمن فئة "السيئ" ظل مستقرًا عند مستويات منخفضة، مما يشير إلى تحسن عام في جودة الهواء خلال السنوات الأخيرة.



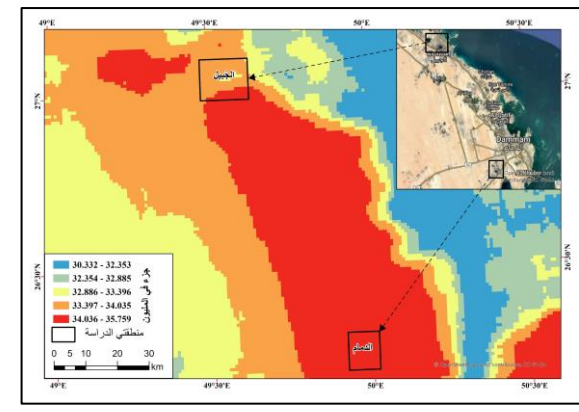
(ب)



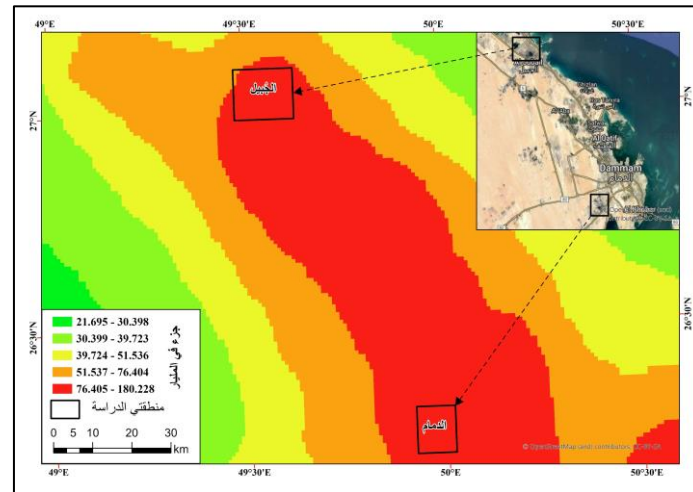
(ا)



(د)



(ج)



(هـ)

شكل ٤. التوزيع المكاني للملوثات الجوية بمنطقة الدراسة للفترة (يناير ٢٠١٩ - يونيو ٢٠٢٤)، (أ) الجسيمات العالقة $PM_{2.5}$ ، (ب) الأوزون O_3 ، (ج) أول أكسيد الكربون CO ، (د) ثاني أكسيد الكبريت SO_2 ، (هـ) ثاني أكسيد النيتروجين NO_2 .

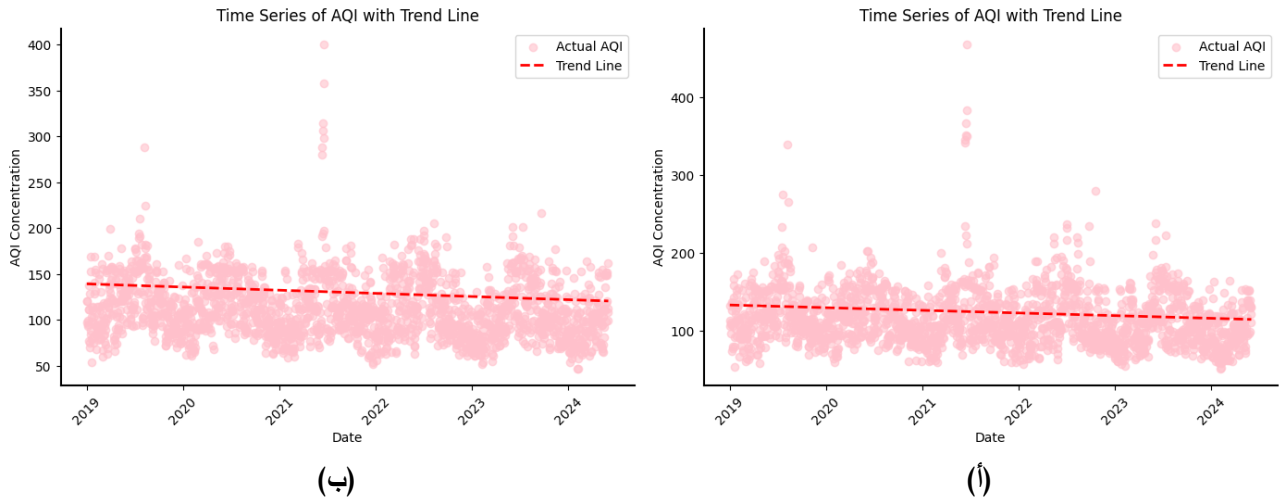
جدول ٤. إحصاءات مؤشر جودة الهواء طبقاً لفترة الدراسة (٢٠١٩-٢٠٢٤) في الدمام والجبيل.

السنة	المدينة	جيد (Good)	مُرْضي (Satisfactory)	معتدل (Moderate)	سيء (Poor)	سيء للغاية (Very Poor)
2019	الدمام	0	118	244	3	0
	الجبيل	0	104	254	6	1
2020	الدمام	0	122	244	0	0
	الجبيل	0	114	250	2	0
2021	الدمام	0	146	212	3	3
	الجبيل	0	118	237	3	6
2022	الدمام	0	156	208	1	0
	الجبيل	0	139	217	9	0
2023	الدمام	0	176	186	3	0
	الجبيل	0	153	209	3	0
2024	الدمام	2	88	62	0	0
	الجبيل	0	91	61	0	0

يُظهر جدول (٥) وشكل (٥) تحليل اتجاه مؤشر جودة الهواء AQI في مدينتي الجبيل والدمام باستخدام انحدار Theil-Sen. توضح القيم السالبة لمعامل الميل (Slope) لكل من الجبيل (-٠,٠٠٩٣٢٣) والدمام (-٠,٠٠٩٤٤١) أن هناك اتجاهًا تنازليًا لمؤشر جودة الهواء في كلا المدينتين خلال فترة الدراسة. يشير هذا الاتجاه إلى تحسن تدريجي في جودة الهواء، حيث يعكس انخفاض قيمة AQI تراجع مستويات التلوث الجوي. ومع ذلك، فإن الفروق بين المدينتين ضئيلة، مما يشير إلى تشابه الأنماط البيئية وتدابير الحد من التلوث في كل منهما. أما القيم الثابتة (Intercept)، والتي تعكس القيمة التقديرية لمؤشر AQI عند بداية فترة الدراسة، فتبلغ ٧٠٠٥,٣٠ للجبيل و ٧٠٩٧,٤١ للدمام، مما يشير إلى أن مستويات التلوث عند نقطة البداية كانت مرتفعة نسبيًا في كلتا المدينتين. تُظهر قيم P-value المنخفضة جدًا (0.0001) أن النتائج ذات دلالة إحصائية قوية، مما يؤكد أن الاتجاه التنازلي لمؤشر AQI ليس ناتجًا عن الصدفة، بل يمثل نمطًا حقيقيًا في جودة الهواء على المدى الطويل.

جدول ٥. اتجاه مؤشر جودة الهواء بالجبيل والدمام باستخدام انحدار Theil-Sen.

المدن	مؤشر جودة الهواء AQI		
	Slope	Intercept	P-value
الجبيل	-0.009323	7005.302169	0.0001
الدمام	-0.009441	7097.407998	0.0001



شكل ٥. اتجاه مؤشر جودة الهواء للفترة من (٢٠١٩-٢٠٢٤)، (أ) الجبيل، (ب) الدمام.

٥-٥ استخدام نموذج XGBoost للتنبؤ المستقبلي بمؤشر جودة الهواء في الجبيل والدمام للفترة (٢٠٢٥-٢٠٣٠)

يهدف هذا التحليل إلى تقييم أداء نموذج XGBoost في التنبؤ بجودة الهواء في كل من الجبيل والدمام، وذلك من خلال مجموعة من الإعدادات التجريبية ومؤشرات الأداء الإحصائي. يعتمد النموذج على ضبط دقيق للمعاملات باستخدام التحقق المتقاطع (Cross-Validation) لضمان تحسين دقة التنبؤ وتقليل الخطأ التنبؤي.

تم تنفيذ التدريب باستخدام خمس طيات (5-Fold Cross-Validation) لضمان استقرار النموذج عبر بيانات مختلفة، بالإضافة إلى اختبار عشرة مرشحين (10 Candidates) لضبط المعلمات المثلى. بلغ إجمالي عمليات التدريب خمسين تكرارًا (50 Iterations) لكل مدينة، مما يشير إلى عملية تدريب مكثفة تهدف إلى تحقيق أقصى أداء ممكن للنموذج. كما في جدول (٦).

فيما يتعلق بمعلمات النموذج، فقد تم ضبط القيم الأساسية وفقًا للآتي:

- عدد الأشجار (n_estimators): تم استخدام 300 شجرة لتعزيز القدرة التنبؤية للنموذج.
- أقصى عمق للأشجار (max_depth): تم تحديد العمق بـ 4 مستويات لتجنب التعقيد الزائد الذي قد يؤدي إلى فرط التكيف مع بيانات التدريب.
- معدل التعلم (learning_rate): تم استخدام 0.1 كمعدل تعلم متوسط لتحقيق التوازن بين سرعة التقارب ودقة النتائج.

- معدل أخذ العينات العشوائي (subsample): تم تحديده عند 0.8 مما يعني أن كل نموذج تدريبي يستخدم ٨٠% من البيانات المتاحة، مما يساعد على تقليل احتمالية التكيف المفرط (Overfitting).
- معدل أخذ الميزات في كل شجرة (colsample_bytree): تم ضبطه على 1.0 مما يعني استخدام جميع الميزات عند بناء كل شجرة قرار.

لقياس دقة النموذج، تم الاعتماد على متوسط الخطأ التربيعي (MSE) ومعامل التحديد (R^2)، حيث تعكس هذه المؤشرات مدى كفاءة النموذج في التنبؤ بقيم جودة الهواء (AQI). أظهرت النتائج أن النموذج حقق أداءً جيدًا في كلتا المدينتين، إلا أن الأداء في الدمام كان أكثر دقة مقارنةً بالجبيل. بالإضافة إلى ذلك تم عمل تحليل المخلفات البيانية (Residuals).

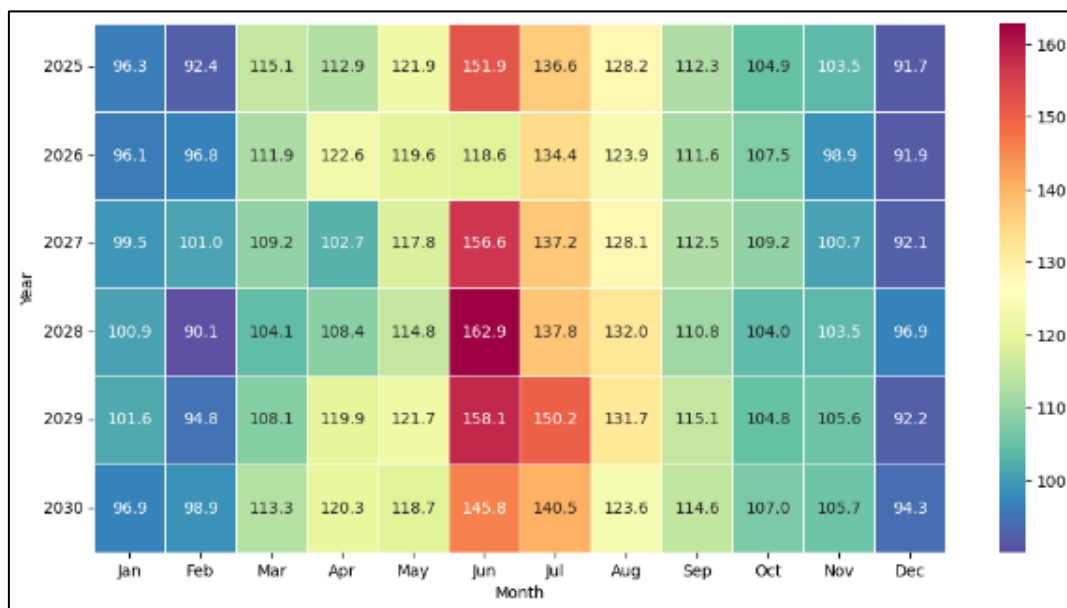
بلغ متوسط الخطأ التربيعي (MSE) 11.51 في الجبيل، في حين كان ٤,١٣ في الدمام، مما يشير إلى أن الأخطاء التنبؤية في الجبيل كانت أكبر من مثيلاتها في الدمام. حقق معامل التحديد R^2 قيمًا مرتفعة في كلتا المدينتين، حيث بلغ ٠,٩٦ في الجبيل و ٠,٩٨ في الدمام، مما يشير إلى أن النموذج يفسر ٩٦% و ٩٨% من التباين في البيانات على التوالي.

تشير نتائج هذا التحليل إلى أن نموذج XGBoost أظهر كفاءة عالية في التنبؤ بجودة الهواء، خاصة في الدمام حيث كانت دقة التنبؤ أكبر وانخفضت قيم الخطأ التربيعي المتوسط. بشكل عام، تؤكد النتائج أن XGBoost نموذج قوي وفعال في التنبؤ بجودة الهواء، مع تحقيق دقة تنبؤية عالية تتجاوز ٩٥%، مما يجعله أداة مناسبة لرصد وتوقع التغيرات في جودة الهواء وتحليل العوامل المؤثرة عليها.

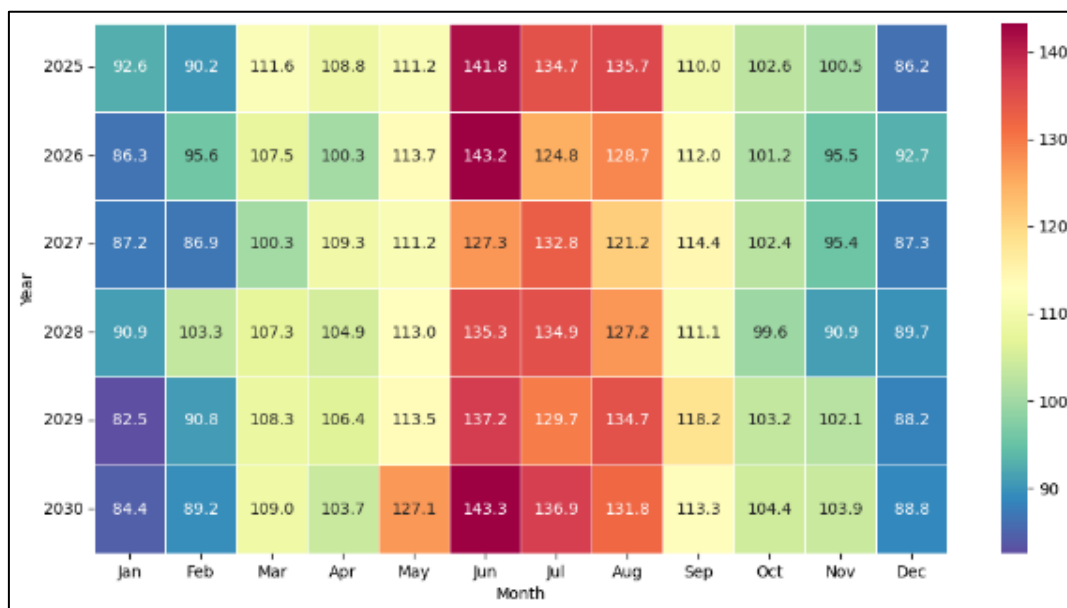
جدول ٦. نتائج نموذج XGBoost للتنبؤ المستقبلي بجودة الهواء في كل من الجبيل والدمام.

الوصف	الجبيل	الدمام
عدد مرات التدريب (Folds)	5	5
عدد التجارب (Candidates)	10	10
إجمالي مرات التدريب	50	50
subsample	0.8	0.8
n_estimators	300	300
max_depth	4	4
learning_rate	0.1	0.1
colsample_bytree	1.0	1.0
متوسط الخطأ التربيعي (MSE)	11.51	4.13
معامل التحديد (R^2)	0.96	0.98

تعرض الأشكال (٦-أ)، و(٦-ب) القيم المتنبئ بها لمؤشر جودة الهواء (AQI) في مدينتي الجبيل والدمام للفترة الممتدة من ٢٠٢٥ إلى ٢٠٣٠. تُستخدم هذه الخرائط الحرارية لتوضيح التغيرات الزمنية والموسمية لجودة الهواء، حيث تعكس الألوان الدافئة (الأحمر والبرتقالي) ارتفاع قيم AQI، مما يشير إلى تدهور جودة الهواء، بينما تمثل الألوان الباردة (الأزرق والأخضر) انخفاض القيم، مما يدل على تحسن جودة الهواء.



(أ)



(ب)

شكل ٦. قيم مؤشر جودة الهواء المتنبئ بها للفترة (٢٠٢٥-٢٠٣٠)، (أ) الجبيل، (ب) الدمام.

في مدينة الجبيل، يظهر نمط موسمي واضح في تغيرات جودة الهواء، حيث تصل قيم AQI إلى أعلى مستوياتها خلال أشهر الصيف (يونيو - أغسطس)، متجاوزة 150 في بعض السنوات. هذا الارتفاع يعكس زيادة تركيز الملوثات الهوائية نتيجة ارتفاع درجات الحرارة وقلّة سرعة الرياح، مما يؤدي إلى تراكم الجسيمات العالقة والملوثات الكيميائية. في المقابل، تنخفض القيم خلال الشتاء (يناير - فبراير)، حيث تتراوح بين 90 و100، مما يشير إلى تحسن جودة الهواء خلال هذه الأشهر، بسبب زيادة معدلات التهوية الطبيعية وانخفاض الأنشطة الصناعية المكثفة.

أما في مدينة الدمام، فيظهر اتجاه موسمي مشابه للجبيل، ولكن بقيم AQI أقل نسبياً. خلال فصل الصيف، ترتفع القيم إلى مستويات تفوق 140، مما يشير إلى تأثيرات بيئية مماثلة لما يحدث في الجبيل. ومع ذلك، تبقى القيم أقل من نظيرتها في الجبيل، مما قد يعكس تفاوتاً في كثافة المصادر الصناعية أو العوامل المناخية المحلية. في فصل الشتاء، تنخفض القيم إلى أقل من 90، مما يعكس جودة هواء أفضل مقارنة بالصيف.

يمكن تفسير هذه الأنماط الموسمية بناءً على عدة عوامل بيئية ومناخية. يؤدي ارتفاع درجات الحرارة خلال الصيف إلى زيادة التفاعلات الكيميائية للملوثات، مما يساهم في تكوين الأوزون الأرضي (O_3) والجسيمات الدقيقة ($PM_{2.5}$) و(PM_{10})، التي تؤثر سلباً على جودة الهواء. كما أن انخفاض سرعة الرياح خلال الصيف يؤدي إلى تراكم الملوثات وعدم تشتتها بفعالية. إضافة إلى ذلك، تلعب العواصف الرملية دوراً مهماً في رفع مستويات AQI خلال الصيف، بينما تساهم زيادة الأمطار والانخفاض العام في النشاط الصناعي خلال الشتاء في تحسين جودة الهواء.

التوقعات المستقبلية لهذه القيم تشير إلى الحاجة إلى تنفيذ تدابير بيئية صارمة للحد من التلوث، خصوصاً خلال أشهر الصيف. يمكن تحقيق ذلك من خلال تعزيز سياسات الحد من الانبعاثات الصناعية، وزيادة المساحات الخضراء، وتحفيز استخدام مصادر الطاقة النظيفة. كما ينبغي تعزيز برامج مراقبة جودة الهواء عبر محطات رصد متطورة لضمان تقييم دقيق لحالة التلوث واتخاذ الإجراءات المناسبة. إضافة إلى ذلك، يجب توعية السكان حول تأثيرات تدهور جودة الهواء والتوصية بتقليل الأنشطة الخارجية خلال فترات الذروة في AQI، خصوصاً في الصيف.

تُظهر هذه التوقعات أن جودة الهواء في الجبيل والدمام ستواجه تحديات بيئية، خاصة خلال فصل الصيف، مع استمرار التفاوت بين المدينتين لصالح الدمام. لذلك، من الضروري اتخاذ إجراءات فعالة للحد من التلوث وتعزيز الاستدامة البيئية، مما يساهم في تحسين جودة الحياة في المنطقة.

يقدم جدول (٧) ملخصاً إحصائياً لتحليل المخلفات البيانية (Residuals) لنموذج XGBoost في مدينتي الدمام والجبيل، ويوضح مدى دقة النموذج في التنبؤ بقيم مؤشر جودة الهواء (AQI) في كل مدينة. يتضح من الجدول أن عدد القيم المستخدمة في التحليل متساوٍ في كلتا المدينتين (١٣ قيمة لكل مدينة)، مما يعزز من موثوقية المقارنة.

تشير القيم الإحصائية إلى أن متوسط المخلفات في الدمام بلغ ٠,٧٨، وهو أعلى من المتوسط في الجبيل (٠,٣٨)، ما يدل على أن الأخطاء التنبؤية في الدمام تميل لأن تكون موجبة أكثر من الجبيل. أما الانحراف المعياري للمخلفات، فقد كان أقل في الدمام (١,٩٥) مقارنة بالجبيل (٣,٥)، ما يعكس تشتتاً أكبر للأخطاء في الجبيل، ويشير إلى أن دقة النموذج في الدمام كانت أفضل وأكثر استقراراً.

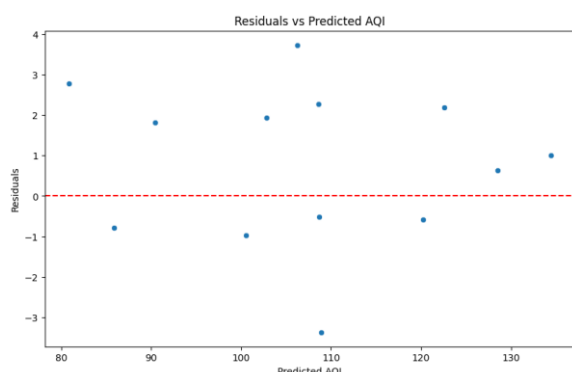
بالنسبة للقيم القصوى، يظهر الحد الأدنى للمخلفات في الجبيل (-٧,١) وهو أقل بكثير من الدمام (-٣,٣٧)، مما يدل على وجود بعض الأخطاء الكبيرة في التنبؤ في الجبيل. كما أن الحد الأقصى للمخلفات في الجبيل (٤,٥) أعلى من الدمام (٣,٧٣)، ما يشير إلى وجود تباين أكبر في نتائج التنبؤ في الجبيل. أما القيم الربعية (الربع الأول، الوسيط، الربع الثالث) فتوضح أن معظم المخلفات في الدمام تتركز حول القيم الموجبة، بينما تتوزع المخلفات في الجبيل بين القيم السالبة والموجبة بشكل أوسع.

بشكل عام، يعكس جدول (٧) أن نموذج XGBoost حقق أداءً تنبؤياً جيداً في المدينتين، مع دقة أعلى وتوزيع أخطاء أضيق في الدمام مقارنة بالجبيل، وهو ما يمكن ربطه بتعقيد العوامل البيئية والصناعية في الجبيل التي تؤثر على جودة الهواء وتجعل التنبؤ أكثر تحدياً. توضح هذه النتائج أهمية الاستمرار في تحسين النماذج التحليلية، خاصة في المناطق الصناعية ذات التغيرات البيئية الكبيرة.

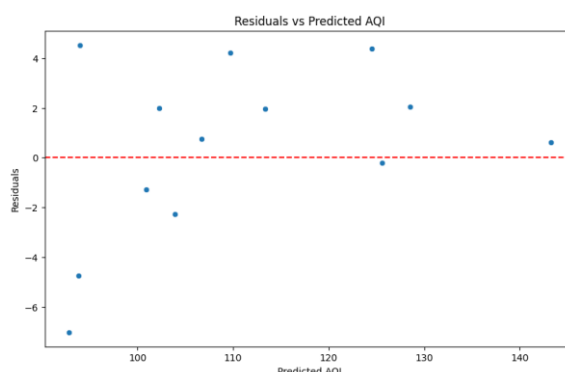
أظهر تحليل المخلفات للنموذج المدرب على بيانات جودة الهواء نتائج إيجابية بشكل عام، مع بعض الملاحظات التي يمكن البناء عليها لتحسين أداء النموذج مستقبلاً. في الشكل (٧) تم رسم المخلفات Residuals مقابل القيم المتوقعة لمؤشر جودة الهواء Predicted AQI. يلاحظ أن المخلفات موزعة عشوائياً حول خط الصفر الأحمر، دون وجود نمط أو اتجاه واضح. هذا التوزيع العشوائي يشير إلى أن النموذج لا يعاني من تحيزات منهجية في التنبؤات، وأن الأخطاء لا ترتبط بقيم AQI المتوقعة. مثل هذا السلوك يُعد دليلاً على أن النموذج يلتقط العلاقات الرئيسية في البيانات بشكل جيد، ولا يوجد ميل واضح لزيادة أو نقصان الأخطاء مع تغير القيم المتوقعة.

جدول ٧. النتائج الإحصائية لتحليل المخلفات البيانية Residuals لنموذج XGBoost في مدينة الدمام والجبلي.

القيمة	الدمام	الجبيل
عدد القيم	13	13
المتوسط	0.78	0.38
الانحراف المعياري	1.95	3.5
الحد الأدنى	-3.37	-7.1
الربع الأول	-0.58	-1.3
الوسيط	1.00	0.76
الربع الثالث	2.19	2.1
الحد الأقصى	3.73	4.5



(ب)



(أ)

شكل ٧. قيم المخلفات البيانية Residuals مقابل القيم المتوقعة لمؤشر جودة الهواء (AQI). (أ) الجبيل، (ب) الدمام.

٦ - المناقشة

تُظهر النتائج التي تم تحليلها اختلافات واضحة في تراكيزات الملوثات الهوائية بين مدينتي الجبيل والدمام، مما يعكس تأثير الأنشطة الصناعية، انبعاثات المركبات، والعوامل الجوية على جودة الهواء في كلتا المدينتين. عند مقارنة هذه النتائج بالدراسات السابقة، يمكن استخلاص بعض التفسيرات والملاحظات المهمة.

أظهرت النتائج ارتفاع متوسط تراكيزات $PM_{2.5}$ في الجبيل (٤٥,٢٨ ميكروغرام/م^٣) مقارنة بالدمام (٤٢,٠٤ ميكروغرام/م^٣)، مع تسجيل قيم قصوى أعلى في الجبيل، مما يتوافق مع تقارير منظمة الصحة العالمية حول تلوث الهواء في المدن الصناعية. أما CO، فقد كانت تراكيزاته متقاربة في الجبيل (٣,٣٧ جزء في المليون) والدمام (٣,٥٢ جزء في المليون)، مما يعكس استقرار الانبعاثات الناجمة عن احتراق الوقود.

وفيما يتعلق بـ NO_2 ، سجلت الجبيل تراكيزات أعلى (١٣,٧٧ جزء في البليون) مقارنة بالدمام (١٢,٤١ جزء في البليون)، ما يشير إلى تأثير الأنشطة الصناعية وحركة المرور المكثفة. كذلك، كانت مستويات SO_2 في الجبيل (٣٢,٣٨ جزء في البليون) أعلى من الدمام (٢٦,٤٨ جزء في البليون)، مع قيم قصوى مرتفعة تعكس التأثير الكبير للصناعات النفطية والطاقة (العيالني، ٢٠١٦).

خُصت هذه النتائج إلى أن الجبيل تشهد عمومًا مستويات أعلى من بعض الملوثات الهوائية مقارنة بالدمام، خاصة $PM_{2.5}$ ، NO_2 ، و SO_2 ، مما يعكس التأثير الكبير للأنشطة الصناعية في الجبيل. تتوافق هذه النتائج مع الدراسات السابقة التي أشارت إلى أن المناطق الصناعية تشهد تركيزات أعلى من هذه الملوثات. بالإضافة إلى ذلك، تشير النتائج إلى أن انبعاثات CO و O_3 مستقرة وقابلة للمقارنة في كلتا المدينتين، مما يعكس مصادر الانبعاثات المتشابهة والعمليات الكيميائية الضوئية المتشابهة.

تعكس الاتجاهات الزمنية للملوثات الجوية في الجبيل والدمام تأثير السياسات البيئية والتغيرات الصناعية على جودة الهواء. أظهرت النتائج انخفاضًا ملحوظًا في تركيز $PM_{2.5}$ بمرور الوقت، وهو ما يتماشى مع الدراسات التي تؤكد دور اللوائح البيئية في تقليل الانبعاثات الصناعية والمرورية (Li, 2020).

كما سجل SO_2 انخفاضًا معنويًا في المدينتين، مما يشير إلى نجاح الجهود في تقليل انبعاثات المصانع ومحطات الطاقة، وهو ما يتفق مع تقارير منظمة الصحة العالمية التي تربط سياسات الحد من الكبريت بتحسين جودة الهواء (WHO, 2021).

على الجانب الآخر، ارتفع NO_2 بشكل ملحوظ في الدمام، مما يعكس زيادة الانبعاثات من حركة المرور والتوسع العمراني، وهي ظاهرة سبق توثيقها في دراسات حول تأثير النمو الحضري على جودة الهواء (Zheng et al., 2018). كما سجل CO ارتفاعًا طفيفًا في الدمام، بينما كان غير دالٍ إحصائيًا في الجبيل، مما قد يشير إلى اختلاف مصادر الانبعاثات في المدينتين. أما O_3 فقد شهد زيادة معنوية في كلتا المدينتين، وهو ما يرتبط بالتفاعلات الكيميائية الضوئية الناجمة عن زيادة NO_2 والمركبات العضوية المتطايرة (VOCs)، كما أشارت دراسات سابقة إلى تزايد الأوزون في المناطق الحضرية نتيجة هذه التفاعلات (Sicard et al., 2020).

يُظهر التوزيع المكاني لتركيزات الملوثات الجوية في مدينتي الجبيل والدمام خلال الفترة من يناير ٢٠١٩ إلى يونيو ٢٠٢٤ تباينًا واضحًا في مستويات التلوث بين المنطقتين، وهو ما يعكس التأثيرات المختلفة للأنشطة الصناعية والمرورية على جودة الهواء.

يُلاحظ أن الفئة "المُرضية" كانت الأكثر شيوعًا في كلا المدينتين، حيث زادت في الجبيل من ١٠٤ أيام في ٢٠١٩ إلى ١٥٣ يومًا في ٢٠٢٣، وفي الدمام من ١١٨ إلى ١٧٦ يومًا خلال نفس الفترة. يُعزى هذا التحسن إلى السياسات البيئية التي تهدف إلى الحد من الانبعاثات الصناعية والمرورية.

انخفض عدد الأيام المصنفة ضمن الفئات "المعتدلة" في كلتا المدينتين، مما يشير إلى تحول نحو جودة هواء أفضل. كما تراجع عدد الأيام ضمن الفئتين "السيئ" و "السيئ جدًا"، حيث لم تسجل الدمام أي أيام ضمن

الفئة "السيئ جدًا"، بينما اختفت هذه الفئة تمامًا من الجبيل بعد عام ٢٠٢١. يتماشى هذا الاتجاه مع التحسن العالمي في جودة الهواء في بعض المناطق الحضرية نتيجة تنفيذ استراتيجيات الحد من الانبعاثات الصناعية والتخطيط العمراني المستدام.

يؤكد تحليل Theil-Sen الاتجاه التنازلي في مؤشر جودة الهواء، حيث بلغت القيم السالبة لمعامل الميل في الجبيل (-0.009323) وفي الدمام (-0.009441)، مما يشير إلى انخفاض تدريجي في مستويات التلوث. تعكس هذه النتائج نجاح تدابير مراقبة الانبعاثات وتقنيات الحد من التلوث مثل استخدام الوقود النظيف وتحسين كفاءة المصانع، وهي نتائج تتوافق مع تقارير منظمة الصحة العالمية (WHO, 2021) حول تأثير التشريعات البيئية في تقليل تلوث الهواء.

يُظهر نموذج XGBoost كفاءة عالية في التنبؤ بجودة الهواء في الجبيل والدمام، حيث حقق R^2 مرتفعًا في كلتا المدينتين (٠,٩٦ و ٠,٩٨ على التوالي)، مما يشير إلى قدرته على تفسير معظم التباين في البيانات. ومع ذلك، لوحظت فروقات في دقة التنبؤ، حيث كانت الأخطاء التنبؤية أعلى في الجبيل ($MSE = 11.51$) مقارنة بالدمام ($MSE = 4.13$)، مما يعكس تأثير الجبيل بعوامل بيئية وصناعية أكثر تعقيدًا.

٧- الخاتمة

أظهرت أظهرت نتائج الدراسة أن جودة الهواء في الجبيل والدمام تتأثر بشكل واضح بالأنشطة الصناعية والمرورية، حيث سجلت الجبيل متوسطات أعلى لبعض الملوثات مثل $PM_{2.5}$ (٤٥.٢٨ ميكروغرام/م^٣)، NO_2 (١٣,٧٧ جزء في البليون) و SO_2 (٣٢,٣٨ جزء في البليون) مقارنة بالدمام ($PM_{2.5} = 42.04$ ميكروغرام/م^٣)، $NO_2 = 12.41$ جزء في البليون، $SO_2 = 26.48$ جزء في البليون، مما يعكس التأثير الكبير للصناعات البتروكيمياوية وكثافة العمليات الصناعية في الجبيل. رغم ذلك، سجلت البيانات الزمنية انخفاضًا ملحوظًا في تركيز $PM_{2.5}$ و SO_2 في كلا المدينتين، حيث بلغ معامل الانحدار لمؤشر $PM_{2.5}$ -0.003885 في الجبيل و-٠,٠٠٤١٨٢ في الدمام، كما انخفض SO_2 بمعدل -٠,٠٠٢٢٠٨ في الجبيل و-٠,٠٠٥٨٧٧ في الدمام سنويًا، وهو ما يعكس نجاح السياسات البيئية في تقليل الانبعاثات الصناعية. في المقابل، لوحظ ارتفاع في تركيز NO_2 و O_3 ، خاصة في الدمام، نتيجة زيادة حركة المرور والتفاعلات الكيميائية الضوئية، مما يستدعي مزيدًا من الإجراءات للحد من هذه الزيادات. كما أظهر تحليل تصنيف جودة الهواء تراجع عدد الأيام المصنفة ضمن الفئات "السيئ" و"السيئ جدًا"، حيث لم تسجل الدمام أي أيام "سيئ جدًا" بعد عام ٢٠٢١، وانخفضت هذه الفئة في الجبيل إلى الصفر في السنوات الأخيرة، مما يدل على تحسن بيئي ملموس. أثبت نموذج XGBoost فعاليته العالية في التنبؤ بجودة

الهواء، حيث بلغت قيمة R^2 في الدمام ٠,٩٨ وفي الجبيل ٠,٩٦، مع متوسط خطأ تربيعي (MSE) بلغ ٤,١٣ في الدمام و ١١,٥١ في الجبيل، مما يعكس قدرة النموذج على تفسير معظم التباين في البيانات ودعم اتخاذ القرار البيئي المبني على البيانات. تعزز هذه النتائج أهمية التخطيط البيئي المستدام، وتوسيع استخدام الطاقة النظيفة، وتكثيف برامج التشجير الحضري، بما ينعكس إيجاباً على جودة الهواء وصحة السكان.

التوصيات

- **تعزيز السياسات البيئية:** ضرورة الاستمرار في تطبيق وتطوير اللوائح البيئية الصارمة لمراقبة الانبعاثات الصناعية والمرورية، مع فرض حدود أكثر تشدداً على المصادر الرئيسية للملوثات، خاصة في المناطق الصناعية مثل الجبيل، بهدف خفض متوسط تركيز $PM_{2.5}$ إلى أقل من ٣٥ ميكروغرام/م³ خلال السنوات الخمس القادمة.
- **توسيع استخدام الطاقة النظيفة:** تشجيع التحول إلى مصادر طاقة نظيفة مثل الغاز الطبيعي والطاقة المتجددة في القطاعات الصناعية والنقل، بما يساهم في خفض انبعاثات NO_2 و SO_2 بنسبة لا تقل عن ٢٠% مقارنة بمستويات ٢٠١٩.
- **تحسين البنية التحتية للنقل المستدام:** توسيع شبكات النقل العام وتعزيز استخدام السيارات الكهربائية والوقود النظيف، خاصة في المدن ذات النمو الحضري السريع مثل الدمام، بهدف خفض متوسط تركيز NO_2 بنسبة ١٠% خلال ثلاث سنوات.
- **تعزيز برامج التشجير والتخطيط العمراني المستدام:** زيادة المساحات الخضراء داخل المدن وحول المناطق الصناعية، حيث تشير الدراسات إلى إمكانية خفض تركيز $PM_{2.5}$ و O_3 بنسبة تصل إلى ١٥% عند مضاعفة مساحة الغطاء النباتي الحضري.
- **استخدام تقنيات الذكاء الاصطناعي لمراقبة جودة الهواء:** تطوير واعتماد نماذج تنبؤية متقدمة مثل XGBoost لمراقبة وتحليل جودة الهواء في الوقت الفعلي، ما يساعد في رصد التغيرات بدقة واستباق فترات التدهور البيئي.
- **زيادة التوعية البيئية:** إطلاق حملات توعية للسكان والعاملين في القطاع الصناعي حول مخاطر الانبعاثات وأهمية الالتزام بالإجراءات البيئية، مع التركيز على الفئات الأكثر تعرضاً لمخاطر التلوث.

- تعزيز التعاون المؤسسي: تشجيع الشراكات بين الجهات الحكومية، القطاع الصناعي، والمؤسسات الأكاديمية لتطوير حلول مستدامة للحد من التلوث الجوي، مع تبادل البيانات والخبرات لتقييم فعالية السياسات البيئية بشكل دوري.

المراجع

أولاً: المراجع العربية

- العلياني، سعيد سعد زاهر، رمضان، أشرف عزمي، والسلطاني، ياس عباس (٢٠١٦). تقييم جودة الهواء في مدينة الجبيل الصناعية بالمملكة العربية السعودية (رسالة ماجستير غير منشورة). جامعة الخليج العربي، المنامة. مسترجع من <http://search.mandumah.com/Record/1014006>

ثانياً: المراجع الأجنبية

- Ahmed, A. A. M., Jui, S. J. J., Sharma, E., Ahmed, M. H., Raj, N., & Bose, A. (2024). An advanced deep learning predictive model for air quality index forecasting with remote satellite-derived hydro-climatological variables. *Science of The Total Environment*, 906, 167234.
- AlGhamdi, A. (2020). Saudi Arabia energy report. *King Abdullah Pet. Stud. Res. Cent*, 19, 1-28.
- Almaliki, A. H., Derdour, A., & Ali, E. (2023). Air Quality Index (AQI) Prediction in Holy Makkah based on machine learning methods. *Sustainability*, 15(17), 13168.
- Al-Rawabdeh, A., Alzu'bi, F., & Almagbile, A. (2025). Quantitative assessment of spatiotemporal variability in air quality within the Amman-Zarqa urban Area, Jordan. *The Egyptian Journal of Remote Sensing and Space Sciences*, 28(3), 395-410.
- Al-Taai, S. H. H., & Mohammed al-Dulaimi, W. A. (2022). Air Pollution: A Study of Its Concept, Causes, Sources and Effects. *Asian Journal of Water, Environment and Pollution*, 19(1), 17-22.
- Banerjee, B., Kundu, S., Kanchan, R., & Mohanta, A. (2024). Examining the relationship between atmospheric pollutants and meteorological factors in Asansol city, West Bengal, India, using statistical modelling. *Environmental Science and Pollution Research*, 1-18.
- Baruah, S., & Mazumder, D. H. (2025). A Review on Application of Machine Learning Techniques coupled with E-Nose in Healthcare, Agriculture and Allied Domains. *IEEE Transactions on Instrumentation and Measurement*.
- Bayhan, K., Başakın, E. E., Gençoğlu, S., Ekmekcioğlu, Ö., & Pham, Q. B. (2025). Utilization of artificial intelligence techniques in predicting air quality index. In *Air Pollution, Air Quality, and Climate Change* (pp. 217-230). Elsevier.
- Chen, G., Chen, S., Li, D., & Chen, C. (2025). A hybrid deep learning air pollution prediction approach based on neighborhood selection and spatio-temporal attention. *Scientific Reports*, 15(1), 3685.
- Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).
- Ghasempour, F., Sekertekin, A., & Kutoglu, S. H. (2021). Google Earth Engine based spatio-temporal analysis of air pollutants before and during the first wave COVID-19 outbreak over Turkey via remote sensing. *Journal of Cleaner Production*, 319, 128599.

- **Ghimire, S., Deo, R. C., Jiang, N., Ahmed, A. M., Prasad, S. S., Casillas-Pérez, D., ... & Yaseen, Z. M.** (2025). Explainable deep learning hybrid modeling framework for total suspended particles concentrations prediction. *Atmospheric Environment*, 121079.
- **Haque, M. N., Sharif, M. S., Rudra, R. R., Mahi, M. M., Uddin, M. J., & Abd Ellah, R. G.** (2022). Analyzing the spatio-temporal directions of air pollutants for the initial wave of COVID-19 epidemic over Bangladesh: Application of satellite imageries and Google Earth Engine. *Remote Sensing Applications: Society and Environment*, 28, 100862.
- **Kahia, M., Omri, A., & Jarraya, B.** (2021). Green energy, economic growth and environmental quality nexus in Saudi Arabia. *Sustainability*, 13(3), 1264.
- **Kaleeswari, P., Ramalakshmi, R., Arun Prasath, T., Muthukumar, A., Kottaimalai, R., & Thanga Raj, M.** (2025). DABiG: Breath pattern classification using the hybrid deep learning with optimal feature selection. *Technology and Health Care*, 09287329241303368.
- **Kaveh, M., Mesgari, M. S., & Kaveh, M.** (2025). A Novel Evolutionary Deep Learning Approach for PM_{2.5} Prediction Using Remote Sensing and Spatial–Temporal Data: A Case Study of Tehran.
- **Khreis, H., Johnson, J., Jack, K., Dadashova, B., & Park, E. S.** (2022). Evaluating the performance of low-cost air quality monitors in Dallas, Texas. *International journal of environmental research and public health*, 19(3), 1647.
- **Kumar, B. R.** (2022). Case: 10 Jubail II industrial city. In *Project finance: Structuring, valuation and risk management for major projects* (pp. 145-150). Springer.
- **Li, Y.** (2020). Comprehensive Speciation Analysis of PM_{2.5}: Quantitative Source Apportionment and Qualitative Characterization of Toxic Organics by Ultra-High Resolution Mass Spectrometry Hong Kong University of Science and Technology (Hong Kong)].
- **Loyola, D. G., Gimeno García, S., Lutz, R., Argyrouli, A., Romahn, F., Spurr, R. J., Pedergrana, M., Doicu, A., Molina García, V., & Schüssler, O.** (2018). The operational cloud retrieval algorithms from TROPOMI on board Sentinel-5 Precursor. *Atmospheric Measurement Techniques*, 11(1), 409-427.
- **Makky, N., & Eghrari, Z.** (2024). Forecasting Atmospheric Air Pollution in Tehran Using Random Forest Model. In *International Conference of Environmental Remote Sensing and GIS* (pp. 57-60).
- **Mohan, S., Kumar, B., & Nejadhashemi, A. P.** (2025). Integration of Machine Learning and Remote Sensing for Water Quality Monitoring and Prediction: A Review. *Sustainability*, 17(3), 998.
- **Mujabar, S., & Rao, V.** (2018). Estimation and analysis of land surface temperature of Jubail Industrial City, Saudi Arabia, by using remote sensing and GIS technologies. *Arabian Journal of Geosciences*, 11, 1-13.
- **Mukundan, A., Huang, C.-C., Men, T.-C., Lin, F.-C., & Wang, H.-C.** (2022). Air pollution detection using a novel snap-shot hyperspectral imaging technique. *Sensors*, 22(16), 6231.
- **Ng, C. F. S., Hashizume, M., Obase, Y., Doi, M., Tamura, K., Tomari, S., Kawano, T., Fukushima, C., Matsuse, H., & Chung, Y.** (2019). Associations of chemical composition and sources of PM_{2.5} with lung function of severe asthmatic adults in a low air pollution environment of urban Nagasaki, Japan. *Environmental pollution*, 252, 599-606.
- **Reddy, P. D., & Parvathy, L. R.** (2022a). Prediction analysis using random forest algorithms to forecast the air pollution level in a particular location. *2022 3rd international conference on smart electronics and communication (ICOSEC)*.
- **Reddy, P. D., & Parvathy, L. R.** (2022b). Prediction analysis using random forest algorithms to forecast the air pollution level in a particular location. In *2022 3rd international conference on smart electronics and communication (ICOSEC)* (pp. 1585-1589). IEEE.
- **Sen, P. K.** (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American statistical association*, 63(324), 1379-1389.
- **Sicard, P., De Marco, A., Agathokleous, E., Feng, Z., Xu, X., Paoletti, E., Rodriguez, J. J. D., & Calatayud, V.** (2020). Amplified ozone pollution in cities during the COVID-19 lockdown. *Science of The Total Environment*, 735, 139542.

- **Stratoulas, D., Jang, B., & Nuthammachot, N.** (2025). Evaluation of urban PM_{2.5} concentrations over 73 major cities and their association with satellite Aerosol Optical Depth: A global analysis of ambient air pollution. *Atmospheric Pollution Research*, 102556.
- **Sujatha, G., Vasantha Kumari, P., & Vanathi, S.** (2025). Machine learning approaches for air quality prediction– A review. In *International Conference on Digital Transformation (AC-ICDT’ 20)*.
- **U.S. Environmental Protection Agency.** (2021). Air Quality Index (AQI) basics. <https://www.airnow.gov/aqi/aqi-basics/>
- **Victoriano, J. M., Santos, M. L. C., Vinluan, A. A., & Carpio, J. T.** (2022). Predicting pollution level using random forest: a case study of Marilao River in Bulacan Province, Philippines. *arXiv preprint arXiv:2202.06066*.
- **Vîrghileanu, M., Săvulescu, I., Mihai, B.-A., Nistor, C., & Dobre, R.** (2020). Nitrogen Dioxide (NO₂) Pollution monitoring with Sentinel-5P satellite imagery over Europe during the coronavirus pandemic outbreak. *Remote Sensing*, 12(21), 3575.
- **WHO.** (2021). Global status report on alcohol and health 2018. World Health Organization.
- **Zhang, C., Niu, X., Wu, H., Ding, Z., Chan, K. L., Kim, J., ... & Liu, C.** (2025). Unleashing the potential of geostationary satellite observations in air quality forecasting through artificial intelligence techniques. *Atmospheric Chemistry and Physics*, 25(2), 759-770.
- **Zheng, B., Tong, D., Li, M., Liu, F., Hong, C., Geng, G., Li, H., Li, X., Peng, L., & Qi, J.** (2018). Trends in China's anthropogenic emissions since 2010 as the consequence of clean air actions. *Atmospheric chemistry and physics*, 18(19), 14095-14111.

Applications of Artificial Intelligence and Remote Sensing for Air Quality Analysis in Dammam and Jubail, Saudi Arabia

Jumah bint Ibrahim Ahmed Ain

Department of Geography, College of Humanities and Social Sciences, King Saud University, Riyadh, Saudi Arabia

Jain@ksu.edu.sa

Abstract. Dammam and Jubail face significant environmental and public health challenges due to elevated air pollution resulting from industrial activities and traffic emissions. This study aimed to assess air quality in both cities during the period 2019–2024 by utilizing satellite data (Sentinel-5P and MODIS) in conjunction with artificial intelligence models. The results revealed that Jubail recorded higher concentrations of fine particulate matter (PM_{2.5}), averaging 45.28 µg/m³, compared to Dammam (42.04 µg/m³). Additionally, NO₂ and SO₂ levels were higher in Jubail, reflecting the pronounced impact of industrial activities, while no significant difference was observed in CO concentrations between the two cities. The XGBoost model was employed to predict air quality, achieving high accuracy (R² between 0.96 and 0.98), with prediction errors being greater in Jubail. The study confirmed that industrial emissions and traffic are the primary sources of pollution, although some improvement in certain pollutants was observed due to environmental policies. The study recommends enhancing vehicle efficiency, expanding green spaces, and developing stricter environmental policies to mitigate rising NO₂ and O₃ levels and to achieve a more sustainable environment.

Keyword: Air Pollution, Future Predictions, Spatial Changes, Industrial Areas, Predictive Modeling, Applied Geography.