

توظيف الشبكات العصبية المعادة وتحليل الموجات للتنبؤ
بتلوث الهواء

رسالة تقدمت بها

ختام وليد قادر الزبيدي

إلى

مجلس كلية علوم الحاسوب والرياضيات في جامعة الموصل وهي جزء من

متطلبات نيل شهادة الماجستير

علوم في الإحصاء

بإشراف

المدرس الدكتور

أسامة بشير شكر الحنون

المستخلص

اكتسبت نمذجة جودة الهواء أهمية كبيرة في التنبؤ بتلوث الهواء الجوي، بسبب آثاره السلبية على صحة الإنسان، والحيوان على حد سواء، وعلى البيئة، وتم التطرق إلى آثار تلوث الهواء على الإنسان والحيوان والنبات، وسائر الكائنات الحية على سطح الكرة الأرضية، وكان لابد بدلاً من معالجة تلوث الهواء بعد حدوثه، أن يتم التنبؤ بهذا التلوث للاستعداد له بما يحول دون التأثير بآثاره السيئة على البيئة، وعلى الكائنات الحية على الأرض، ومن بينها الإنسان. وإن دراسة الجسيمات المعلقة في الهواء (Particular matter (PM₁₀)) والتكهن بها ضروري للتقليل والسيطرة على الأضرار البيئية وصحة الإنسان. وقد تعددت طرق التنبؤ بتلوث الهواء حيث تم استخدام نموذج الإنحدار الخطي المتعدد (Multiple linear regression (MLR) كطريقة إحصائية خطية للتنبؤ بمتغير PM₁₀ من خلال تأثره بمتغيرات الأرصاد الجوية المقابلة، لذلك فقد يعكس نتائج غير دقيقة عند استخدامه مع مجموعات البيانات غير الخطية. حيث أن طريقة التراصف الزمني في أنماط مختلفة تم اقتراحها لتحسين تلك النتائج وتحقيق التجانس، ويتضمن مرادفة المواسم المتشابهة في السنوات المختلفة سوية لتكوين متغير جديد مختلف عن الأصلي. لتحسين نتائج التنبؤ تم استخدام الشبكات العصبية المعادة Recurrent neural network (RNN) كأسلوب غير خطي لتستخدم بعد التوفيق مع نموذج MLR ضمن الطريقة الهجينة MLR-RNN. كذلك تم استخدام تحليل المويجات Wavelet للتعامل مع عدم التأكدية واستخدامه في ترشيح النتائج المستحصلة من طريقة MLR-RNN والحصول على تحسين أكثر لنتائج التنبؤ من خلال ما يسمى الطريقة الهجينة RNN-Wavelet واعتماداً على نموذج MLR. إن نتائج التنبؤ بشكل عام كانت الأفضل باستخدام أسلوب التراصف الزمني مع استخدام الطريقة المقترحة الهجينة RNN-Wavelet.

Abstract

Particulate matter (PM_{10}) studying and forecasting is necessary to control and reduce the damage of environment and human health. There are many pollutants as sources of air pollution may effect on PM_{10} variable. This type of dataset can be classified as nonlinear and suffer from the uncertainty. Studied datasets have been taken from climate station in Malaysia. Multiple linear regression (MLR) is used as linear statistical method for PM_{10} forecasting through its influencing by corresponding climate variables, therefore it may reflect inaccurate results when used with nonlinear datasets. Time stratified (TS) method in different styles is proposed for satisfying more homogeneity of datasets. It includes ordering similar seasons in different years together to formulate new variable smoother than their original. To improve the results of forecasting, recurrent neural network (RNN) has been suggested as nonlinear method to be used after combining with MLR in hybrid MLR-RNN. Wavelet analysis is proposed filtering the result of MLR-RNN method for more improving of forecasting results by treating uncertainty problem through RNN-Wavelet hybrid method based on MLR model. In general, the best results of forecasting were for using RNN-Wavelet with TS method.