

استنباط نماذج تحويلية بتقنية الغابة العشوائية وتحليل الانحدار (RFR) وتحليل الانحدار المتعدد (MLR) لتقدير السعة الحقلية ونقطة الذبول الدائم لترب منطقة وادي الحي، بسهل الجفارة، ليبيا.

يونس ضو زايد الزليط¹، فاطمة مختار مصباح ابورزيزة²، يوسف صالح المبروك حامد¹

1- قسم التربة والمياه - كلية الزراعة - جامعة طرابلس - طرابلس - ليبيا

2- مركز البحوث الزراعية - طرابلس - ليبيا

المستخلص

أجريت هذه الدراسة لبناء واختبار قدرة نماذج تحويلية مستنبطة باستخدام أداة الغابات العشوائية وتحليل الانحدار (RFR)، التي تعد أحد نماذج التعلم الآلي، ومقارنتها بنماذج مستنبطة باستخدام تحليل الانحدار الخطي المتعدد (MLR)، وذلك لتقدير خاصيتي السعة الحقلية ونقطة الذبول الدائم لبيانات عينات سطحية من (157) قطاعاً ممثلاً لتربة وادي الحي بسهل الجفارة، والتي تم الحصول عليها من دراسة Hidroprojeat. وكانت البيانات المدخلة لاستنباط النماذج هي النسب المئوية للرمل والسلت والطين والكثافة الظاهرية والكثافة الفعلية والكربون العضوي، وكانت جودة نماذج MLR المستنبطة مختلفة باختلاف المدخلات في النموذج التحويلي والخاصية المقدر، فكان أفضل نموذج متحصل عليه لتقدير السعة الحقلية بتحليل الانحدار الخطي المتعدد عند استخدام نموذج الرمل والسلت والطين كمدخلات، حيث كانت قيم MEA و RMSE و R² و (C2) 2.32 % و 2.89 % و 0.84 و 0.77 على التوالي، في حين كانت أفضل النتائج المتحصل عليها للتنبؤ بنقطة الذبول الدائم للنموذج الذي تم فيه استخدام المدخلات الرمل والسلت والطين والكثافة الظاهرية والكثافة الفعلية والكربون العضوي والسعة الحقلية المقاسة، بقيم للمعايير MEA و RMSE و R² و (C2) 1.78 % و 2.23 % و 0.56 و 0.56 على التوالي. في العموم، كانت النماذج MLR المستنبطة ذات جودة منخفضة ولا يمكن الاعتماد عليها كدوال للتنبؤ بخاصيتي السعة الحقلية ونقطة الذبول الدائم. وعند استخدام RFR لبناء نماذج تحويلية للتنبؤ بكلا الخاصيتين، تم الحصول على نتائج أفضل مما تم الحصول عليها من تحليل الانحدار الخطي المتعدد؛ حيث كانت المعايير الإحصائية للنماذج التحويلية المستنبطة للسعة الحقلية أفضل بكثير، بحيث تراوحت قيمة MEA و RMSE و R² ومعامل ناش (C2) لنماذج تقدير السعة الحقلية، (0.88-1.15 %) و (1.14-1.24 %) و (0.93-0.93) و (0.99)، كما كانت القدرة التقديرية للنماذج المستنبطة لتقدير نقطة الذبول الدائم أفضل في جميع النماذج، حيث تراوحت قيمة MEA و RMSE و R² ومعامل ناش (C2) ما بين 0.61-0.88 % و (0.68-1.43 %) و (0.93-0.93) و (0.99)، بالرغم من أن جميع النماذج بالمدخلات المختلفة كانت جيدة، إلا أن تقديرات النماذج تحسنت كلما زاد عدد المدخلات فيها. وللتأكد من عمل هذه النماذج في بيئة نظم المعلومات الجغرافية تم الاعتماد على أفضل نموذج مستنبط باستخدام RFR لإنتاج خرائط تنبؤ مكاني بطريقة مقلوب المسافة الموزون IDW، وتم الحصول على خرائط مشابهة ومقاربة لخرائط التنبؤ المكاني للقيم المقاسة، وبذلك توصي هذه الدراسة بزيادة استخدام واختبار النماذج المختلفة للتعلم الآلي، ومن أهمها الغابة العشوائية وتحليل الانحدار لاشتقاق نماذج لخصائص قد يتعذر قياسها. الكلمات الدالة: الدوال التحويلية، الوادي الحي، الغابة العشوائية وتحليل الانحدار، الترب الليبية.

المقدمة

غالبًا ما يتم قياس كل من السعة الحقلية ونقطة الذبول الدائم للتربة باستخدام جهاز أوعية الضغط، الذي يتطلب أخذ عينات تربة من الحقل ونقلها إلى المعمل، مما يجعل عمليات القياس مكلفة وتحتاج إلى عمالة بخبرات عالية، وتتطلب هذه العملية -أيضًا- وقتًا طويلًا لإتمامها يصل أحيانًا إلى عدة أشهر (Hillel, 2004). كما أن عدد العينات المأخوذة من الحقل يظل صغيراً مقارنة بمساحة الحقل (Pachepsky *et al.*, 2015)، الأمر الذي دعا العديد من الباحثين إلى التفكير في تقدير قيم السعة الحقلية ونقطة الذبول من خلال ربطها بعلاقات مع بعض قيم خواص التربة السهلة والسريعة القياس، واشتقاق ما يعرف بالدوال التحويلية للتربة (Pedotransfer function)، حيث يتم استغلال العلاقات المتبادلة بين بيانات الخواص الطبيعية والكيميائية للتربة لتقدير المفقود منها. على الرغم من نجاح هذه الدوال في تقدير خصائص التربة غير المعلومة بالاعتماد على بيانات خواص التربة المتاحة، إلا أن هذه الدوال تظل محصورة الاستخدام في المناطق التي تم استنباط هذه الدوال من بياناتها (عاشور، 2017). الأمر الذي دعا العديد من الباحثين حول العالم لاشتقاق دوال محلية باستخدام وسائل اشتقاق مختلفة، ابتداءً من تحليل الانحدار الخطي وغير الخطي المتعدد، ووصولاً إلى الشبكات الاصطناعية من أجل الحصول على نماذج موثوقة لتقديرات خواص التربة المختلفة (Ezlit *et al.*, 2014)، إلا أن هناك قصوراً وتحديات تؤثر سلباً في دقة النماذج المستنبطة. فقد راجع (Pachepsky *et al.*, 2015) التطورات الأخيرة في أساليب وتقييمات الدوال التحويلية للتربة حول العالم، وتم تحديد الثغرات ونقاط الضعف في الدوال التحويلية المقترحة في البحوث، وقاموا بتلخيصها في عدة نقاط، وهي: أولاً- الحاجة إلى تكبير نطاق الدوال التحويلية لتصبح قابلة للتطبيق في مناطق مختلفة حول العالم. ثانياً- الحاجة إلى المزيد من الدوال التحويلية للتربة الحدية، مثل: الترب المالحة والترب

الجيرية والترب الجبسية والترب الطينية المحتوية على معادن الطين من مجموعة السمكتايت المتمددة. ثالثاً- الحاجة إلى متغيرات تحكم في العمليات البيو جيوكيميائية، مثل: نماذج تطور الكربون والنيتروجين في التربة (Eihermueller *et al.*, 2013). إن من أهم أسباب تباين تقدير الدوال التحويلية للتربة هو استخدام بيانات مجمعة من ترب بمساحات كبيرة، ولا يمثل عدد نقاط البيانات المجمعة من هذه المساحات التغيرات في خواص التربة المختلفة، نظراً للتباين الكبير بين نقاط أخذ القياسات، وبالتالي لا يتم تمثيل بيانات خواص التربة بشكل صحيح.

يمكن إطلاق مصطلح الدوال التحويلية المكانية (GeoPedoTransfer Function) أو اختصار (GPTF) على النموذج الذي يتم فيه دمج علاقات خواص التربة عند موضع محدد مع التقديرات المكانية لتغير الخاصية المراد قياسها (Kang *et al.*, 2020). وقد قدم مؤخراً عدد من الباحثين محاولات لاشتقاق GPTF في بيئة نظم المعلومات الجغرافية باستخدام الذكاء الاصطناعي (Artificial Intelligent) أو اختصار (AI) الذي يحوي العديد من النماذج المرنة والقابلة للتعلم، ويمكن استخدامها تحت بيئة نظم المعلومات الجغرافية، (Tuncay *et al.*, 2017).

تُعدّ أساليب وتقنيات ML كثيرة ومتعددة الاستخدامات، بحيث لا يمكن الإحاطة بكافة الأساليب المستخدمة نظراً للتقدم المتسارع في هذا المجال، إلا أن هناك العديد من هذه التقنيات التي تم استخدامها لبناء نماذج تحويلية دقيقة للتربة، ومن أهم هذه الأساليب الغابة العشوائية وتحليل الانحدار (Random Forest and Regression)، أو اختصاراً (RFR)، حيث استخدمها كل من (Beguin *et al.*, 2017) و (Khanal *et al.*, 2018) في التنبؤ بخواص التربة المختلفة، واستخدمها كل من (Brungard *et al.*, 2015) و (Hounkpatin *et al.*, 2018) في نموذج تحويلي لتصنيف الترب المختلفة بناءً على بيانات الخواص

(2017; Tuncay *et al.*, 2020). لذلك هدف هذا البحث إلى استنباط واختبار مرونة نماذج تحويلية جيومكانية بالاعتماد على أحد تقنيات الذكاء الاصطناعي ألا وهي الغابة العشوائية وتحليل الانحدار (RFR)، وذلك لتقدير كل من السعة الحقلية ونقطة الذبول الدائم في بيئة نظم المعلومات الجغرافية، ومقارنتها بنماذج تم اشتقاقها لنفس البيانات من منطقة وادي الحي في سهل الجفارة، باستخدام تقنية تحليل الانحدار الخطي المتعدد (MLR).

المواد وطرائق البحث

الوصف العام لمنطقة الدراسة

تقع منطقة الدراسة في وادي الحي (سابقا الوادي الميت)، الواقع في جنوب وسط سهل الجفارة شمال غرب ليبيا، وبالتحديد ما بين خطي طول $12^{\circ} 21'$ إلى $13^{\circ} 02'$ شرقاً، ودائرتي عرض $32^{\circ} 26'$ إلى $32^{\circ} 32'$ شمالاً (شكل 1)، وتبلغ مساحة منطقة الدراسة حوالي 11000 هكتار، حيث تنتشر في منطقة الدراسة العديد من أنواع ترب المناطق الجافة وشبه الجافة، كذلك تبين سيادة وانتشار الترب حديثة التكوين الرملية ذات النظام الرطوبي الجاف (Hidroprojekat, 1974).

مصادر البيانات ومعالجتها

تم الحصول على بيانات التربة المستخدمة في هذه الدراسة من دراسة (Hidroprojekat 1974) التي تمت خلال فترة السبعينات مع بداية دراسات الترب الليبية، وقد اعتمدت هذه الدراسة في تصنيف حبيبات التربة (Hidroprojekat 1974)؛ حيث كان تصنيف حبيبات التربة كما هو معتمد من الجمعية العالمية لعلوم التربة (ISSS, 1927)، وأن نتائج أو تحديد النسب المختلفة لحبيبات التربة تم باستخدام طريقة الماصة. تم الحصول على خريطة مواقع قطاعات التربة ورقياً والتي أعدتها شركة Hidroprojekat، وتم إجراء المسح الضوئي لعدد 23 لوحة. تضم هذه اللوحات عدد 157 قطاع، وتم بعد ذلك تحويلها إلى صورة رقمية، ثم إرجاعها جغرافياً، ورسم خريطة رقمية لموقع الدراسة

المختلفة للتربة، واستخدمها (Were *et al.*, 2015) و (Emadi *et al.*, 2020) في تخطيط الكربون العضوي، واستخدمها (Feng *et al.*, 2019) في تقدير درجة حرارة التربة بالاعتماد على البيانات المناخية للمنطقة، وطور بها (Bondre 2020) نموذج للتنبؤ بالإنتاجية بناءً على عمليات التسميد.

من ناحية أخرى تم استخدام طرائق التنبؤ المكاني لدراسة التبع المكاني والزمني لخصائص التربة الطبيعية والكيميائية، التي تحدد القيمة بناءً على موقعها بالنسبة لبيانات نقاط معلومة القيمة تحيط بالنقطة المراد تقدير الخاصية عندها، حيث استخدمت طرائق مختلفة، منها: طريقة مقلوب المسافة الوزنية (IDW) وطريقة كريكينج (Kriging). قدم العديد من الباحثين تطبيقات للإحصاء المكاني للتنبؤ بخواص التربة المختلفة مكانياً ومن أمثلتها (Laslett *et al.*, 1987) و (Sigua and Hudnall 2008) و (Hosseini *et al.*, 2014) و (Gia Pham *et al.*, 2019). وقام (Hojati *et al.*, 2023) بالتنبؤ بالتوزيع المكاني ورسم الخرائط الرقمية لتركيز الفوسفور المتاح في التربة في شمال محافظة خوزستان في إيران، حيث استخدم ثلاثة نماذج للتعليم الآلي، وهي Random Forest (RF) و Artificial Neural Network (ANN) و Support Vector Regression (SVR)، لتطوير العلاقة التنبؤية لتقدير تركيز الفوسفور في سطح التربة فحسب النتائج المتحصل عليها تم اختيار خوارزمية RF كأفضل خيار. إلا أنه وبالرغم من كفاءة أسلوب التنبؤ المكاني في تقدير وتخطيط خواص التربة المختلفة، إلا أن درجة دقته اعتمدت بشكل كبير على مدى تجانس التوزيع المكاني لخواص التربة المختلفة، ولا يأخذ هذا الأسلوب التغيرات في الخواص عند النقاط المختلفة (انبيهة، 2017).

إن دمج أسلوب الاشتقاق للدوال التحويلية مع تخطيط الخواص وربطها مكانياً في نموذج تحويلي يعتمد على استخدام أساليب الذكاء الاصطناعي في التقدير يمكن أن يزيد من دقة تقديرات الخواص الطبيعية (Kang *et al.*

والكثافة الفعلية (2.56-2.75 جرام/سم³)، والسعة الحقلية (14.41-35.72%)، ونقطة الذبول الدائم (2.7-18%)، والكربون العضوي (0.06-1.68 %) لكل موقع؛ حيث تم استخدام البيانات التي تم تجميعها في اشتقاق الدوال التحويلية لتقدير السعة الحقلية ونقطة الذبول الدائم للنقاط المعلومة بها كافة الخصائص المدروسة، وتم إعداد جميع البيانات المستخدمة بصيغة CSV ليتسنى استخدامها في برامج نظم المعلومات الجغرافية.

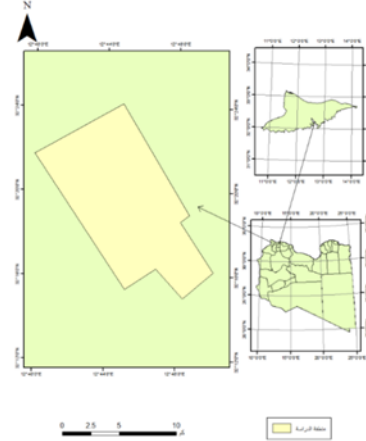
اشتقاق نماذج التربة التحويلية باستخدام الانحدار المتعدد

تم تحليل البيانات باستخدام برنامج Excel 10، بحيث تم دراسة علاقة كل من السعة الحقلية ونقطة الذبول بالخواص الطبيعية للتربة، بافتراض وجود علاقة خطية بين هذه الخصائص وكل من السعة الحقلية ونقطة الذبول الدائم، وحيث أن الحقيقة الفيزيائية أنه في حالة قيم المدخلات في النموذج مساوية للصفر فإن قيم كل من السعة الحقلية ونقطة الذبول تساوي صفر، لذلك تم تحليل الانحدار بقيمة ثابتة للجزء المقطوع من المحور الصادي مساوية للصفر. تم ربط السعة الحقلية بالنسب المختلفة لحبيبات الرمل والست والطين أولاً، ومن ثم تم إضافة المتغيرات المساعدة بداية بالكثافة الظاهرية والكثافة الفعلية ونسبة المادة العضوية، ممثلة في الكربون العضوي، بناءً على عدد المتغيرات المستقلة في النموذج، وقد أمكن الحصول على أربع نماذج خطية بمدخلات متعددة. أما في حالة نقطة الذبول فتم استخدام نفس الأسلوب مع إضافة نموذج خامس يتضمن قيم السعة الحقلية التجريبية من أجل الحصول على تقديرات أكثر دقة.

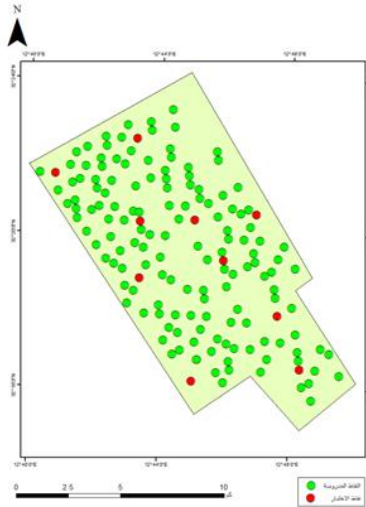
استنباط نماذج التربة التحويلية باستخدام التعليم الآلي بتقنية الغابة العشوائية وتحليل الانحدار (RFR)

تم استخدام أسلوب الغابة العشوائية وتحليل الانحدار (RFR)، وهو أحد أساليب التعليم الآلي التي يمكن الاستفادة منها في بناء دوال تحويلية جيوميكانية،

موزع عليها مواقع القطاعات التي استخدمت بياناتها معرفة بالإحداثيات العالمية المترية (شكل 2)، وذلك باستخدام برنامج Arc pro3.1. وتجدر الملاحظة أن الحدود المعتمدة في هذه الدراسة هي الحدود الخارجية للقطاعات التي تم توقيعهها في الخريطة الرقمية وليست الحدود الإدارية التي من الممكن أن تتوفر لدى الجهات الرسمية بالدولة الليبية.



شكل 1. الموقع العام لمنطقة الدراسة.



شكل 2. توزيع نقاط أخذ العينات في منطقة الدراسة.

تم إعداد وتجهيز قاعدة بيانات للطبقة السطحية من قطاعات التربة المجمعة في هذه الدراسة باستخدام Excel sheets V. 10. وتمثلت هذه البيانات في العمق السطحي للعينة داخل القطاع (أقل من 30 سم)، ونسب الرمل والست والطين والتي تراوحت (60.7-97.9%)، والكثافة الظاهرية (1.44-1.79 جرام/سم³)،

$$C^2 = 1 - \left(\frac{F}{F_0} \right) \quad (3)$$

$$F = \sum_{i=1}^i (P_i - O_i)^2 \quad (4)$$

$$F_0 = \sum_{i=1}^i (P_i - \bar{O}_i)^2 \quad (5)$$

حيث O^- متوسط القيم البيانات المقاسة و O القيم المقاسة و P قيم البيانات المقدرة، وهو قيمة من غير وحدات، ويكون أداء النموذج أفضل كلما اقتربت القيمة المحسوبة له من الواحد، وتشير القيمة الموجبة بأن النموذج مقبول، إلا إن القيم التي تساوي صفر والقيم السالبة تشير إلى ضعف قدرة التقدير للنموذج. كما تم استخدام قيم R^2 الناتج من تحليل الانحدار الخطي البسيط بين القيم التجريبية والقيم المقدرة التي تم حسابها بالنماذج المختبرة لقيمة كل من السعة الحقلية ونقطة الذبول للتقييم، وذلك بافتراض أن الجزء المقطوع من المحور الصادي يساوي صفر في معادلة الخط المستقيم، بحيث تصبح العلاقة بين القيم المحسوبة بواسطة كل نموذج مختبر مع القيم التجريبية كما يلي:

$$P = \alpha O \quad (6)$$

حيث تمثل α قيمة ميل العلاقة بين المقدرة والتي تشير إلى ملائمة النموذج بشكل عام، والقيم التجريبية أو المقاسة، إن قيمة معامل التحديد (R^2) المحسوبة تشير إلى مدى تقارب القيم المقاسة مع القيم التجريبية. فكلما اقتربت القيمة من الواحد ذل ذلك على جودة تقدير النموذج المختبر.

تم استخدام أفضل النماذج المستنبطة بواسطة أسلوب RFR في إنتاج خريطين لكل من السعة الحقلية ونقطة الذبول الدائم للعمق السطحي للتربة لكامل منطقة الدراسة، حيث تم استخدام طريقة مقلوب المسافة الموزون IDW في إنتاج هذه الخرائط نظرا لعدم انطباق شروط استخدام طريقة kriging لإجراء التحليل المكاني على بيانات كل من السعة الحقلية ونقطة الذبول الدائم (انبيه، 2017). تمت مقارنة خرائط السعة الحقلية ونقطة الذبول المنتجة من

من خلال الأداة المضمنة في برنامج ArcGIS Pro 3.1 لبناء نماذج تحويلية لترب منطقة الدراسة. وبالرغم من قدرة هذه الأداة في التعامل مع البيانات المكانية وتضمينها في بناء النماذج التحويلية للتربة، إلا أن الاختبار تم تحديده فقط في قدرتها على نمذجة البيانات غير المكانية، لتسهيل عملية المقارنة مع النماذج التحويلية التي تم اشتقاقها من خلال الانحدار الخطي المتعدد لبيانات منطقة الدراسة، واختبار مدى قدرتها على إنتاج نماذج تحويلية قادرة على التنبؤ بكل من السعة الحقلية ونقطة الذبول في منطقة الدراسة.

تقييم النماذج المستنبطة

لتقييم النماذج التحويلية للتربة التي تم الحصول عليها سواء من تحليل MLR أو باستخدام RFR، تم استخدام عدد من المؤشرات الإحصائية التي تقيم مدى جودة النماذج المستنبطة ما بين القيم المقدرة بواسطة النماذج التحويلية للتربة والقيم الفعلية أو التجريبية، وهذه المؤشرات هي:

1 - متوسط الخطأ المطلق (MAE)، ويمكن حسابه باستخدام المعادلة التالية:

$$MAE = \frac{1}{n} \sum_{i=1}^n |P_i - O_i| \quad (1)$$

2- الجذر التربيعي لمربع الخطأ (RMSE)، ويمكن حسابه بالمعادلة التالية:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - O_i)^2} \quad (2)$$

حيث O قيمة البيانات المقاسة و P قيم البيانات المقدرة، ووحداته هي نفس وحدات البيانات، وفي هذه الدراسة هي النسبة المئوية لقيم كل من السعة الحقلية ونقطة الذبول، ويصبح أداء النموذج أفضل كلما اقتربت قيمة (MAE) من الصفر ويضعف أداء النموذج كلما ابتعد.

3 - معامل ناش -ساتكليف (C_2)، ويمكن حسابه بالنموذج التالي:

البيانات المقدرة مع تلك المنتجة من البيانات التجريبية بصرياً. كما تم مقارنة التقارب ما بين المؤشرات الإحصائية للتحليل المكاني IDW لخرائط كل من السعة الحقلية ونقطة الذبول كمؤشر على دقة تقدير نماذج RFR.

النتائج والمناقشة

تم استنباط النماذج التحويلية للسعة الحقلية ونقطة الذبول الدائم لمنطقة الدراسة باستخدام كل من تحليل الانحدار الخطي المتعدد MLR وطريقة الغابة العشوائية وتحليل الانحدار RFR كل على حدا وبمدخلات مختلفة، تم الحصول على أربع نماذج تحويلية من كل طريقة حسب عدد المدخلات في هذه النماذج (جدول 1) لتقدير السعة الحقلية. حيث تم

اشتقاق النماذج أولاً باستخدام النسب المئوية للرمل (sand) والسلت (silt) والطين (clay)، ثم نموذج النسب المئوية للرمل والسلت والطين بإضافة الكثافة الظاهرية (pb)، ثم نموذج النسب المختلفة للرمل والسلت والطين والكثافة الظاهرية والكثافة الفعلية (ps)، ثم نموذج النسب المئوية للرمل والسلت والطين والكثافة الظاهرية والكثافة الفعلية والكربون العضوي (C). كما تم اشتقاق نماذج نقطة الذبول الدائم (pwp) بنفس المدخلات للسعة الحقلية مع إضافة نموذج خامس يتضمن مدخلات النسب المئوية للرمل والسلت والطين والكثافة الظاهرية والكثافة الفعلية والكربون العضوي، بالإضافة إلى السعة الحقلية (fc).

جدول 1. مدخلات واختصارات النماذج التحويلية التي تم استنباطها باستخدام كل من تحليل الانحدار المتعدد (MLR) وطريقة الغابة العشوائية وتحليل الانحدار (RFR).

المتغيرات المدخلة			
اسم النموذج	اسم النموذج	المستنبط بطريقة	المستنبط بطريقة
RFR	MLR		
RFRpwp1	RFfc1	Mpwp1	Mfc1
RFRpwp2	RFfc2	Mpwp2	Mfc2
RFRpwp3	RFfc3	Mpwp3	Mfc3
RFRpwp4	RFfc4	Mpwp4	Mfc4
RFRpwp5	-	Mpwp5	-

الدوال التحويلية للسعة الحقلية ونقطة الذبول

الدائم باستخدام تحليل الانحدار الخطي المتعدد

يعرض جدول (2) النماذج المستنبطة لكل من السعة الحقلية ونقطة الذبول الدائم باستخدام تحليل الانحدار المتعدد. وتوضح المؤشرات الإحصائية لنواتج تحليل الانحدار المتعدد مدى تفاوت مقدرة النماذج المستنبطة في تقدير قيم كل السعة الحقلية ونقطة الذبول (جدول 3). وتشير قيم اختبار F أن النماذج الخطية المتعدد هي نماذج ملائمة لوصف العلاقة ما

بين كل من السعة الحقلية ونقطة الذبول الدائم، وبالرغم من أن قيم R^2 و R^2 المعدل كانت عالية، إلا أن أداء هذه النماذج كان ما بين مقبول إلى ضعيف، ويظهر ذلك من خلال قيم كل الخطأ القياسي (St error) الذي تراوح في نماذج تقدير السعة الحقلية ما بين 2.09 إلى 2.52 وفي نماذج السعة الحقلية ما بين 1.8 إلى 1.52، مما يعكس أن هناك تشتت واختلافات كبيرة في التقديرات، في حين كانت معالم النماذج المستنبطة مختلفة من حيث المعنوية من خلال تباين معنوية قيم

t-test، والتي تعكس عدم ثبات النماذج وعدم قابليتها للتعميم على الترب المشابهة في مناطق أخرى.

جدول 2. نماذج تقدير السعة الحقلية ونقطة الذبول الدائم باستخدام تحليل الانحدار المتعدد.

رقم النموذج	نماذج تقدير السعة الحقلية المستنبطة بتحليل الانحدار المتعدد
Mfc1	$fc = 0.13Sand + 0.436Silt + 1.047Clay.$
Mfc2	$fc = 0.644Sand + 0.849Silt + 1.3Clay - 30.365\rho_b.$
Mfc3	$fc = -0.131Sand - 0.013Silt + 0.472Clay - 36.877\rho_b + 33.369\rho_s.$
Mfc4	$fc = -0.124Sand - 0.03Silt + 0.48Clay - 35.827\rho_b + 32.352\rho_s + 0.937c.$
نماذج تقدير نقطة الذبول الدائم المستنبطة بتحليل الانحدار المتعدد	
Mpwp1	$pwp = 0.334Sand + 0.554Silt + 0.603Clay - 17.793\rho_b.$
Mpwp2	$pwp = 0.334Sand + 0.554Silt + 0.603Clay - 17.793\rho_b.$
Mpwp3	$pwp = 0.091Sand + 0.284Silt + 0.343Clay - 19.835\rho_b + 10.462\rho_s.$
Mpwp4	$pwp = 0.102Sand + 0.257Silt + 0.354Clay - 18.1\rho_b + 8.783\rho_s + 1.5480c.$
Mpwp5	$pwp = 0.136Sand + 0.265Silt + 0.223Clay - 8.254\rho_b - 0.108\rho_s + 1.2910c + 0.275fc$

تقدير كل من السعة الحقلية ونقطة الذبول الدائم خلال مرحلتي الاشتقاق والتقييم (جدولي 4 و 5). ففي مرحلة الاشتقاق تجاوزت قيم R^2 لجميع نماذج السعة الحقلية ونقطة الذبول 0.9، وكانت قيم الخطأ القياسي (ST error) منخفضة بشكل كبير في هذه المرحلة، بحيث تراوحت قيمه ما بين 0.015 و 0.02 للنماذج السعة الحقلية و 0.019 و 0.023 للنماذج تقدير نقطة الذبول الدائم. كذلك أظهرت المؤشرات الإحصائية في حالة التقييم الأولي للنماذج على البيانات التي لم تدخل في اشتقاق النموذج (10% من البيانات الكلية) أن نماذج الغابة العشوائية كانت لها موثوقية عالية في التقدير. بشكل عام يمكن ملاحظة أن تقديرات النماذج ذات المدخلات الأكثر كانت أفضل. والذي يتيح للتعلم الآلي باستخدام طريقة الغابة العشوائية وتحليل الانحدار فرصة دراسة سلوك تغير كل من السعة الحقلية ونقطة الذبول مع المتغيرات المختلفة.

استنباط نماذج التربة التحويلية باستخدام التعلم الآلي بتقنية الغابة العشوائية وتحليل الانحدار تم إدخال البيانات المستخدمة في تحليل الانحدار للاستخدام في التنبؤ بطريقة الغابة العشوائية، حيث تم إنشاء طبقة نقاط في برنامج Arc pro3.1، وذلك لاشتقاق نماذج تحويلية لتقدير كل من FC و PWP، وبنفس المدخلات السابق ذكرها (جدول 1)، وتم ترتيب النماذج حسب المدخلات، ثم تم اختيار عدد مئة شجرة بمساهمة كل النقاط ويقوم البرنامج تلقائياً بفصل 10% من البيانات المدخلة لتقييم النماذج، وتم تغيير عدد الأشجار بشكل متكرر، حتى الحصول على قيم تنبؤات جيدة بدلالة المعايير الإحصائية أثناء عملية الاشتقاق فكانت 80 شجرة. تشير النتائج المتحصل عليها خلال مرحلة استنباط النماذج بتقنية التعلم الآلي باستخدام أداة الغابة العشوائية وتحليل الانحدار في بيئة نظم المعلومات الجغرافية إلى أن النماذج المطورة كانت دقيقة في

جدول 3. المؤشرات الإحصائية للنماذج المستنبطة من تحليل الانحدار المتعدد.

R ²	Ad. R ²	St. Error	F-test	t-test							النموذج	
				Sand	Silt	clay	OC	BD	PD	FC		
0.99	0.98	2.52	4479	27.41	4.9	19.3	*	*	*	*	Mfc1	تحليل تحليلي
0.99	0.98	2.49	3450	2.15E-53	5.9E-05	2.5E-41	0.026	*	*	*	Mfc2	
0.99	0.98	2.19	3566	8.16	7.95	20.6	1.54	-6.7	*	*	Mfc3	
0.99	0.99	2.09	3286	0.53	0.82	0.04	0.22	4.33E-13	7.32E-05	*	Mfc4	
0.96	0.96	1.80	1254	9.70	5.09	11.82	*	*	*	*	Mpwp1	نقطة النول العام
0.96	0.96	1.74	1007	7.59	4.01	11.32	3.35	*	*	*	Mpwp2	
0.97	0.96	1.63	928	5.24	6.40	12.38	2.83	-4.80	*	*	Mpwp3	
0.97	0.96	1.62	780	0.59	1.36	2.04	2.70	-5.02	1.45	*	Mpwp4	
0.97	0.97	1.52	766	0.87	1.54	1.36	2.40	-1.93	-0.022	4.68	Mpwp5	

يعرض شكل 3 مقارنة بيانية لنتائج تقدير السعة الحقلية ما بين القيم التجريبية والقيم التقديرية للنماذج المستنبطة، باستخدام تحليل الانحدار المتعدد. حيث يلاحظ ابتعاد القيم التجريبية عن خط التقدير المثالي. وهذا ظهر بوضوح من خلال نتائج المؤشرات الإحصائية المعروضة في جدول 6. حيث تراوحت المؤشرات MAE و RMSE و R2 و (C2) لنماذج تقدير السعة الحقلية، (1.6 - 2.32 %) و (2 - 2.89 %) و (0.66-0.84) و (0.37-0.77) على التوالي، وفي العموم كان أفضل النماذج بناء على هذه المؤشرات النموذج Mfc1 الذي كانت مدخلاته نسب الرمل والسلت والطين. كما يعرض شكل 3 مقارنة بيانية لنتائج تقدير نقطة الذبول الدائم للنماذج المستنبطة باستخدام تحليل الانحدار المتعدد، حيث يظهر أن النموذج Mppw1 قدم أفضل أداء، إلا أن شكل 3 بين ضعف كل النماذج في تقدير نقطة الذبول الدائم. وهذا ما أكدته نتائج المؤشرات الإحصائية، حيث تراوحت مؤشرات كل من MAE و RMSE و R2 و (C2) لنماذج تقدير نقطة الذبول، (1.62-1.78 %) و (1.90-2.23 %) و (0.39-0.56) و (0.35-0.56) على التوالي، مما يعكس عدم ملائمة تحليل الانحدار المتعدد لاشتقاق نماذج دقيقة، خصوصاً لتقدير قيم نقطة الذبول الدائم.

يوضح شكل 5 مقارنة بيانية لنتائج تقدير نماذج السعة الحقلية المستنبطة باستخدام طريقة الغابة العشوائية وتحليل الانحدار، حيث يبدو بوضوح انخفاض التشتت حول التقدير المثالي لقيم السعة الحقلية، مما يؤكد أن تقديرات السعة الحقلية بواسطة هذه النماذج أكبر دقة مقارنة بتحليل الانحدار الخطي المتعدد، وهذا ما تم تأكيده من خلال المؤشرات الإحصائية المستخدمة في تقييم النماذج (جدول 7)، حيث تراوحت المؤشرات MAE و RMSE و R2 ومعامل ناش (C2) لنماذج تقدير السعة الحقلية، (0.88 - 1.15 %) و (1.14-1.24 %) و (0.93-0.95) و (0.99-0.93).

كما يمكن ملاحظة أن نماذج تقديرات السعة الحقلية، لها موثوقية أكبر مقارنة بنقطة الذبول الدائم، إلا أن تقديرات نقطة الذبول الدائم مقبولة من الناحية الإحصائية.

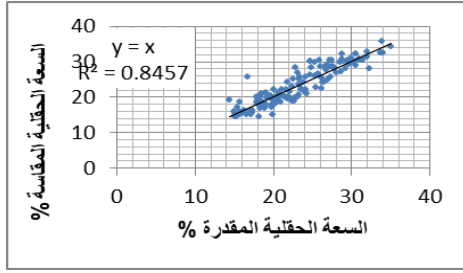
جدول 4. قيم المعايير الإحصائية لمرحلة اشتقاق النماذج التحويلية باستخدام RFR لتقدير FC و PWP.

النماذج	St Error	R ²	
Mfc1	0.02	0.935	السعة الحقلية
Mfc2	0.017	0.955	
Mfc3	0.016	0.96	
Mfc4	0.015	0.966	
Mppw1	0.023	0.902	نقطة الذبول الدائم
Mppw2	0.021	0.928	
Mppw3	0.02	0.929	
Mppw4	0.019	0.941	
Mppw5	0.019	0.944	

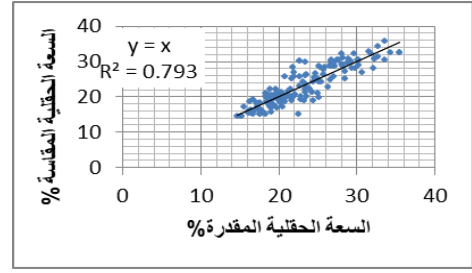
جدول 5. قيم المعايير الإحصائية لمرحلة تقييم النماذج التحويلية باستخدام RFR لتقدير FC و PWP.

النماذج	Standard Error	R2	
Mfc1	0.109	0.771	السعة الحقلية
Mfc2	0.084	0.897	
Mfc3	0.097	0.856	
Mfc4	0.09	0.863	
Mppw1	0.171	0.519	نقطة الذبول الدائم
Mppw2	0.147	0.493	
Mppw3	0.149	0.711	
Mppw4	0.113	0.767	
Mppw5	0.104	0.848	

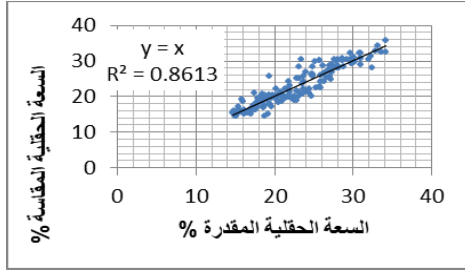
تقييم أداء النماذج المستنبطة بتحليل الانحدار المتعدد وباستخدام الغابة العشوائية وتحليل الانحدار.



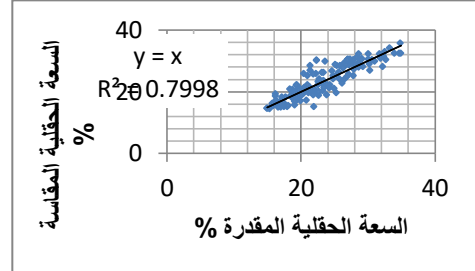
(ج)



(أ)



(د)

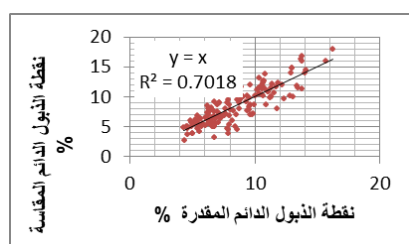


(ب)

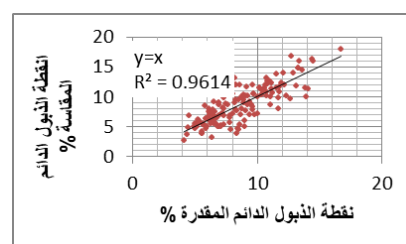
شكل 3. مقارنة القيم المقاسة للسعة الحقيقية والقيم المقدرة للنماذج المستنبطة بطريقة تحليل الانحدار المتعدد. (أ) نموذج (Mfc1)، (ب) نموذج (Mfc2)، (ج) نموذج (Mfc3)، (د) نموذج (Mfc4).

جدول 6. معايير الأداء لمقارنة النماذج المختلفة المستخدمة في تقدير السعة الحقيقية ونقطة الذبول الدائم باستخدام تحليل الانحدار المتعدد.

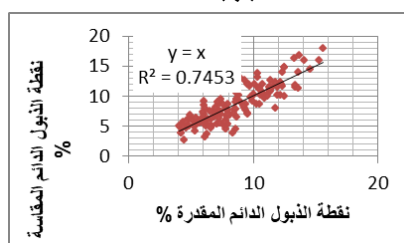
معايير الأداء للنماذج المختلفة المستنبطة لتقدير السعة الحقيقية				
النماذج	C^2	R^2	RMSE (%)	MAE (%)
Mfc1	0.77	0.84	2.00	1.60
Mfc2	0.67	0.73	2.42	1.92
Mfc3	0.53	0.66	2.89	2.43
Mfc4	0.37	0.69	2.73	2.33
معايير الأداء للنماذج المختلفة المستنبطة لتقدير نقطة الذبول الدائم				
النماذج	C^2	R^2	RMSE (%)	MAE (%)
Mpwp1	0.37	0.40	2.25	1.62
Mpwp2	0.36	0.39	2.27	1.73
Mpwp3	0.35	0.39	2.30	1.77
Mpwp4	0.38	0.40	2.23	1.78
Mpwp5	0.56	0.56	1.90	1.50



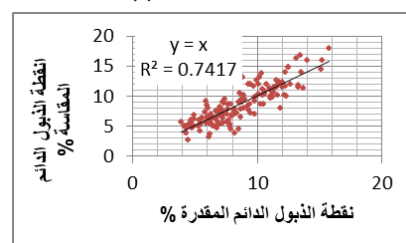
(ب)



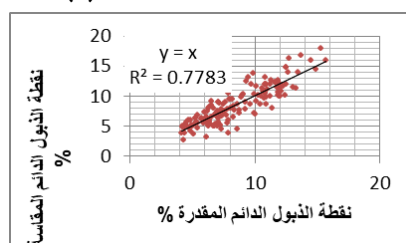
(إ)



(د)



(ج)

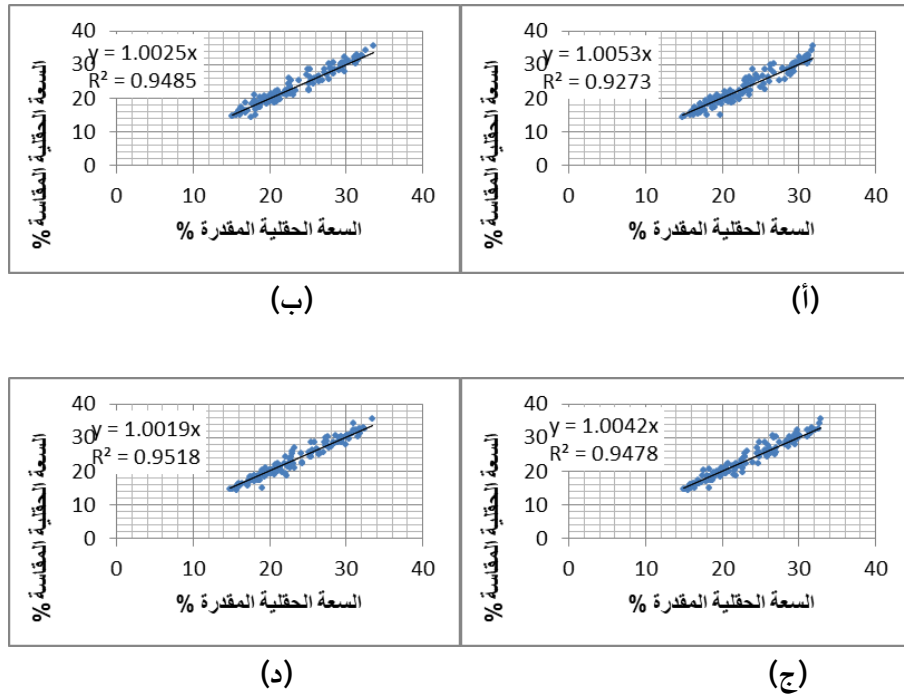


(هـ)

شكل 4. مقارنة القيم المقاسة لنقطة الذبول الدائم والقيم المقدرة للنماذج المستنبطة. (أ) نموذج (Mppw1) ، (ب) نموذج (Mppw2) ، (ج) نموذج (Mppw3) ، (د) نموذج (Mppw4) ، (هـ) نموذج (Mppw5).

جدول 7. معايير الأداء لمقارنة النماذج المختلفة المستخدمة في تقدير السعة الحقلية ونقطة الذبول الدائم باستخدام تحليل الغابة العشوائية.

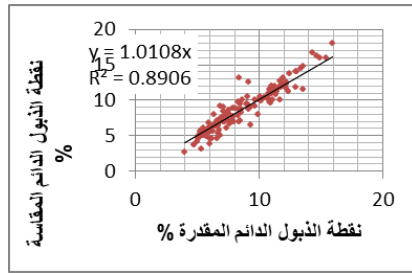
معايير الأداء لمقارنة النماذج المختلفة المستخدمة في تقدير السعة الحقلية باستخدام الغابة العشوائية				
النماذج	C2	R ²	RMSE (%)	MAE (%)
RFRfc1	0.93	0.93	1.14	1.15
RFRfc2	0.95	0.95	1.23	0.95
RFRfc3	0.95	0.95	1.24	0.92
RFRfc4	0.99	0.95	1.19	0.88
معايير الأداء لمقارنة النماذج المختلفة المستخدمة في تقدير نقطة الذبول الدائم باستخدام الغابة العشوائية				
النماذج	C2	R ²	RMSE (%)	MAE (%)
RFRpwp1	0.86	0.86	1.20	0.88
RFRpwp2	0.90	0.89	1.04	0.75
RFRpwp3	0.90	0.90	1.01	0.76
RFRpwp4	0.91	0.91	0.93	0.65
RFRpwp5	0.93	0.94	0.82	0.61



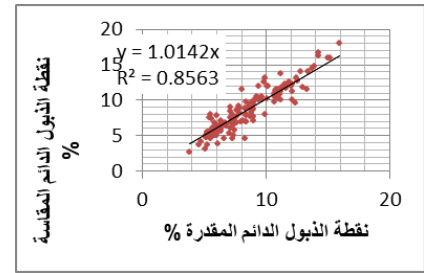
شكل 5. مقارنة القيم المقاسة للسعة الحقلية والقيم المقدرة للنماذج المستنبطة بطريقة الغابة العشوائية وتحليل الانحدار. (أ) نموذج الرمل والسلت والطين (RFRfc1)، (ب) نموذج (RFRfc2)، (ج) نموذج (RFRfc3)، (د) نموذج (RFRfc4).

كما يمكن ملاحظة أن التشتت في المقارنة البيانية ما بين قيم نقطة الذبول الدائم المقدرة التجريبية باستخدام النماذج المستنبطة بطريقة الغابة العشوائية انخفض كثيراً مقارنة بطريقة بالنماذج المستنبطة بواسطة الانحدار المتعدد، حيث ظهر ذلك جلياً في شكل 6. وهذا ما تم تأكيده من خلال قيم المؤشرات الإحصائية (جدول 7). حيث تراوحت المؤشرات MAE و RMSE و R^2 و (C2) لنماذج تقدير نقطة الذبول، (0.61 - 0.88 %) و (0.68 - 1.43 %) و

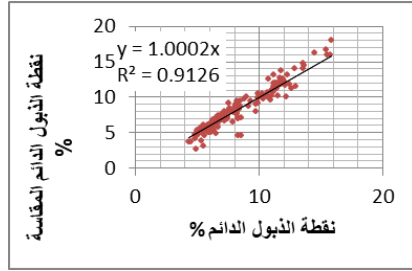
(0.93-0.95) و (0.93-0.99). إن ما يمكن ملاحظته أن طريقة الغابة العشوائية وتحليل الانحدار كانت الأفضل في تقدير السعة الحقلية ونقطة الذبول، وأنه كلما زاد عدد المدخلات في النموذج زادت دقة النموذج. إلا أن جميع النماذج وباختلاف مدخلاتها تعتبر الأفضل مقارنة بالنماذج الخطية التي تم استنباطها بواسطة تحليل الانحدار المتعدد.



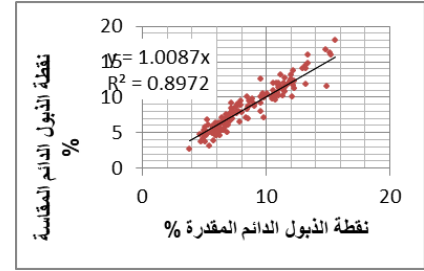
(ب)



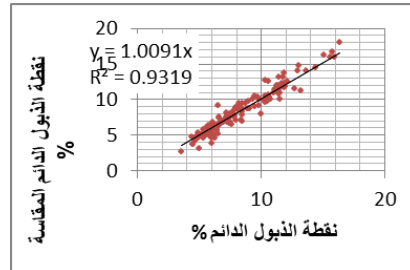
(أ)



(د)



(ج)



(هـ)

شكل 6. مقارنة القيم المقاسة لنقطة الذبول الدائم والقيم المقدرة للنماذج المستنبطة باستخدام الغابة العشوائية وتحليل الانحدار. (أ) نموذج (RFRpwp1)، (ب) نموذج (RFRpwp2)، (ج) نموذج (RFRpwp3)، (د) نموذج (RFRpwp4)، (هـ) نموذج (RFRpwp5).

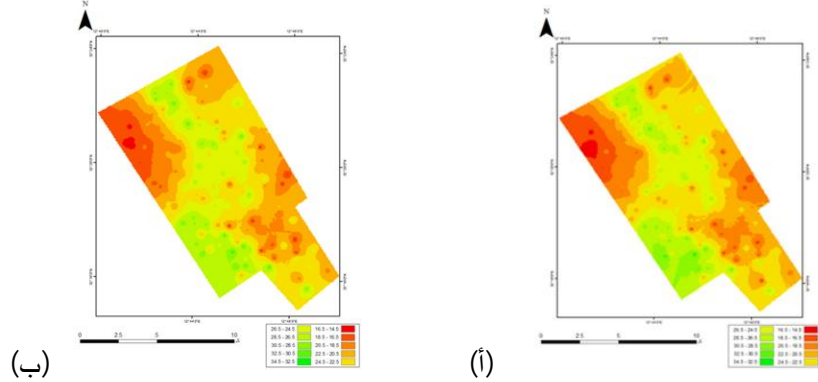
بنفس مقياس الرسم وتدرج القيمة. ويمكن ملاحظة التقارب الكبير بين التوزيع المكاني الناتج من البيانات التجريبية والبيانات المقدرة في كلا الخريطين وللخاصيتين. وبالنظر إلى التقديرات الإحصائية لتقييم التحليل المكاني باستخدام IDW لكلا الخريطين المعروفتين في جدول (8)، يمكن ملاحظة التقارب ما بين هذه القيم في خريطي FC بالبيانات المقدرة والتجريبية، حيث كان ما بين MAE 1.95 و 1.89% و RMSE ما بين 2.33 و 2.29% و C2 ما بين 0.69 و 0.70 للبيانات الحقيقية والمقدرة على التوالي، وكانت قيم R2 متساوي لكلا الخريطين (0.72)، مما يدل على مدى دقة نموذج الغابة العشوائية في تقدير السعة الحقلية. أما بخصوص نقطة الذبول الدائم، فكانت نتائج

خرائط السعة الحقلية ونقطة الذبول التقديرية بأفضل نموذج والبيانات التجريبية

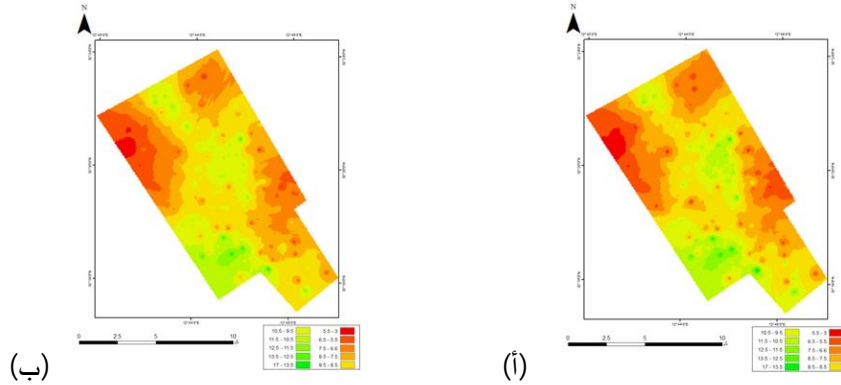
تم استخدام أفضل النماذج المستنبطة باستخدام RFR في إنتاج خريطين لكل من السعة الحقلية (RFRfc4) ونقطة الذبول الدائم (RFRpwp5) للعمق السطحي لكامل منطقة الدراسة، وقد تم استخدام مقلوب المسافة الوزنية لإنتاج هذه الخرائط نظرا لعدم انطباق شروط استخدام طريقة kriging (القماطي، 2020). تم استخدام بيانات FC و PWP التجريبية والمقدرة بنموذج mfc4 و Rfrpwp5 التي تم استنباطها، ويعرض شكلا 7 و 8 الخرائط المنتجة من كل من البيانات التجريبية والبيانات المقدرة من النموذجين المختارين لكل من FC و pwp، بحيث تم إنتاج الخرائط

بين 2.03 و 1.94 % و C2 ما بين 0.49 و 0.53، وهذه القيم المتقاربة للمؤشرات الإحصائية تعكس التقارب الكبير بين نواتج الخريطين.

المؤشرات الإحصائية متقاربة جدا بين نواتج التحليل المكاني باستخدام IDW، حيث كانت قيمتي R2 ما بين 0.57 و 0.62 و MAE ما بين 1.77 و 1.68 % و RMSE



شكل 7. توزيع FC المكاني في منطقة الدراسة بطريقة (IDW)، (أ) القيم التجريبية (ب) القيم المقدرة.



شكل 8. توزيع pwp المكاني في منطقة الدراسة بطريقة (IDW)، (أ) القيم التجريبية (ب) القيم المقدرة.

جدول 8. معايير الأداء لمقارنة طريقة مقلوب المسافة الموزون (IDW) المستخدمة في التنبؤ للسعة الحقلية ونقطة الذبول الدائم باستخدام تحليل الغابة العشوائية.

معايير الأداء لمقارنة طريقة مقلوب المسافة الموزون للسعة الحقلية ونقطة الذبول الدائم.				
النماذج	C ²	R ²	RMSE (%)	MAE (%)
Fc data	0.70	0.72	2.29	1.89
RFRfc4	0.69	0.72	2.33	1.95
Pwp data	0.53	0.62	1.94	1.68
RFRpwp5	0.49	0.57	2.03	1.77

الاستنتاج

التحويلية للتربة وبدقة متفاوتة، إلا إن تقنيات تعلم الآلة توفر حلاً فعالاً لتطوير نماذج تحويلية مرنة، بحيث يمكن ربط خصائص التربة كمتغيرات مستقلة

بالرغم من أن تقنية تحليل الانحدار المتعدد قد تم استخدامها خلال العقود السابقة في استنباط النماذج

Beguín J., Fuglstad, G. A., Mansuy, N. and Paré, D. 2017 Predicting soil properties in the Canadian boreal forest with limited data: Comparison of spatial and non-spatial statistical approaches. *Geoderma*. 306: 195-205.

Bondre, D. A. Prediction of crop yield and fertilizer recommendation using machine learning algorithms. 2020.

Brungard, C. W., Boettinger, J. L., Duniway, M. C., Wills, S. A., and Edwards, Jr. T. C. 2015. Machine learning for predicting soil classes in three semi-arid landscapes. *Geoderma*. 239: 68-83.

Emadi, M., Taghizadeh-Mehrjardi, R., Cherati, A., Danesh, M., Mosavi, A., and Scholten, T. 2020. Predicting and mapping of soil organic carbon using machine learning algorithms in northern Iran. *remote Sensing*. 12(14): 2234.

Feng, Y., Cui, N., Hao, W., Gao, L., and Gong, D. 2019 Estimation of soil temperature from meteorological data using different machine learning models. *Geoderma*, 338, 67-77.

Galton, F 1894. *Natural inheritance* (5th ed). Macmillan and company, New York.

Gia Pham, T., Kappas, M., Van Huynh, C., Hoang Khanh Nguyen, L. 2019. Application of Ordinary Kriging and Regression Kriging Method for Soil Properties Mapping in Hilly Region of Central Vietnam. *ISPRS International Journal of Geo-Information*. 2019, 8, 147.

<https://doi.org/10.3390/ijgi8030147>

Hojati, S., Biswas, A., and Norouzi Masir, M. 2023. Comparing machine learning algorithms

بالخاصية المراد تقديرها، مع إمكانية ربطها –أيضا- ببعض المتغيرات المكانية مثل حالة سطح التربة واستخدام الأراضي في بيئة نظم المعلومات الجغرافية. حيث أثبت أسلوب RFR (أحد تقنيات الذكاء الاصطناعي)، والذي تم استخدامه في هذه الدراسة لاستنباط نماذج تحويلية لتقدير كل من السعة الحقلية ونقطة الذبول فعالية كبيرة في تقديرات القيم التجريبية، مما يفتح الباب أمام بناء نماذج تحويلية مكانية مرنة يمكن تطويرها بناء على توفر البيانات المكانية و اللامكانية حول التربة، بحيث يمكن بناء قاعدة بيانات دقيقة حول خصائص التربة المحلية؛ لذلك توصي الدراسة بالتوسع في استخدام واختبار المزيد من نماذج الذكاء الاصطناعي لتطوير نماذج تحويلية للتربة، وذلك لتقدير خصائص تربة أخرى صعبة القياس أو مكلفة، وينصح باتباع واستخدام نموذج الغابة العشوائية وتحليل الانحدار لبناء قاعدة بيانات وطنية للاستفادة من البيانات المتوفرة من التقارير والدراسات السابقة الكثيرة التي قامت بها العديد من الشركات على كامل التراب الليبي؛ بحيث يمكن توظيف هذه البيانات كمدخلات للنماذج التحويلية للتربة.

المراجع

القماطي، أسماء علي. 2020. تطبيق الاحصاء المكاني في تخطيط بعض الخواص الكيميائية لترب المنطقة الشمالية الشرقية من ليبيا (حالة دراسية: عين حزام، قرية بطة، تاكنس). رسالة ماجستير (غير منشورة) كلية الزراعة جامعة طرابلس

انبيه، أسامة مصطفى. 2017. تأسيس قاعدة بيانات مكانية لبعض الخواص الكيميائية وإنشاء خرائط تفسيرية لترب منطقة سهل الجفارة. رسالة ماجستير (غير منشورة) كلية الزراعة جامعة طرابلس.

عاشور، إيناس محمد. 2017. تقدير بعض خواص التربة المائية باستخدام بيانات التربة سهلة القياس. رسالة ماجستير (غير منشورة) كلية الزراعة جامعة طرابلس.

- prediction methods for soil pH. *Journal of Soil Science*. 38, 325-341.
- Pachepsky, Y. A., Rajkai, K. and Tóth, B. 2015. Pedotransfer in soil physics: trends and outlook- A review. *AgrokémTalanta*. 64(2):339–360.
- Pachepsky, Y. A., Timlin, D., Varallyay, G., 1996. Artificial neural networks to estimate soil water retention from easily measurable data. *Soil Science Society of America Journal*. 60:727-73.
- Hosseini, S. Z. Kappas, M. Bodaghabadi, M. B. Chahouki, M. A. Z. Khojasteh, E. R. 2014. Comparison of Different Geostatistical Methods for Soil Mapping Using Remote Sensing and Environmental Variables in Poshtkouh Rangelands, Iran. *Polish Journal of Environmental Studies*, 23(3), 737-751.
- Sigua, G. S. and Hudnall, W. H. 2008. Kriging analysis of soil properties. *Journal of Soils and Sediments*. 8. 10.1007/s11368-008-0003-7.
- Tuncay, T., Baskan, O., Bayramın I., Dengiz, O. and Kılıç, S. 2017 Geostatistical approach as a tool for estimation of field capacity and permanent wilting point in semi-arid terrestrial ecosystem/ *Archives of Agronomy and Soil Science*.
- Hillel, D. 2004. *Introduction to environmental soil physics*. Academic Press, New York.
- for predicting and digitally mapping surface soil available phosphorous: a case study from southwestern Iran. *Precision Agriculture*, 1-26.
- Houkpatin, K. O., Schmidt, K., Stumpf, F., Forkuor, G., Behrens, T., Scholten, T. and Welp, G. 2018. Predicting reference soil groups using legacy data: A data pruning and random forest approach for tropical environment (Dano catchment, Burkina Faso). *Scientific reports*, 8(1): 1-16.
- International Society of Soil Science. 1927. Conclusions of the first commission meeting at Rothamsted-Harpenden 1926; Conclusions de la conférence de la première commission à Rothamsted-Harpenden 1926; Beschlüsse der Konferenz der ersten Kommission in Rothamsted-Harpenden 1926, p. 16. Brno, Czechoslovakia.
- Kang, Y., Xiaoyan, L., Dehua, M., Zongming, W. and Mingxuan, L. 2020. Combining artificial neural network and ordinary kriging to predict wetland soil organic carbon concentration in China's Liao River Basin.
- Khanal S., Fulton, J., Klopfenstein A., Douridas N., and Shearer, S. 2018. Integration of high resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield. *Computers and Electronics in Agriculture*. 153: 213-225.
- Laslett, G.M., Mc Bratney, A.B., Pahl, Hutchinson, M.F., 1987. Comparison of several spatial

Developing pedotransfer Models Using Random Forest and Regression and Multiple Linear Regression to Estimate Field Capacity and Permanent Wilting Point for Wadi Al-Hai Soils, Libya

Younes Daw Z. Ezlit¹, Fatima Mokhtar Aburzeza², and Yuosef Saleh m. Hamid¹

1- Department of Soil and Water, Faculty of Agriculture, University of Tripoli.

2 – Agriculture Research Center - Tripoli - Libya.

ABSTRACT

This study aimed to develop and evaluate the performance of a transformational predictive model using Random Forest Regression (RFR), a machine learning tool, and compare it with a model derived from Multiple Linear Regression (MLR). The objective was to estimate the field capacity (FC) and permanent wilting point (PWP) of soils using surface samples collected from 157 representative soil profiles in Wadi Al-Hai, Al-Jafara Plain, Libya. These samples were obtained from the Hidroprojekat report (1974). The input variables for model development included the percentages of sand, silt, and clay, as well as bulk density, particle density, and organic carbon content. The performance of the MLR-derived models varied depending on the input variables and the soil property being predicted. The best MLR model for predicting field capacity was obtained when sand, silt, and clay were used as inputs, with MAE, RMSE, R^2 , and Dash (C^2) values of 2.32%, 2.89%, 0.84, and 0.77, respectively. For the prediction of the permanent wilting point, the best MLR model included sand, silt, clay, bulk density, particle density, organic carbon, and measured field capacity as inputs, resulting in MAE, RMSE, R^2 , and Dash (C^2) values of 1.78%, 2.23%, 0.56, and 0.56, respectively. However, the overall predictive performance of the MLR models was inadequate for reliable estimation of field capacity and permanent wilting point. When the RFR method was employed to develop predictive models for both properties, it outperformed the MLR models significantly. The RFR-derived model for field capacity yielded superior results, with MAE, RMSE, R^2 , and Dash (C^2) values ranging from 0.88–1.15%, 1.14–1.24%, 0.93–0.95, and 0.93–0.99, respectively. Similarly, the RFR model demonstrated higher predictive accuracy for the permanent wilting point, achieving MAE, RMSE, R^2 , and Dash (C^2) values between 0.61–0.88%, 0.68–1.43%, 0.93–0.95, and 0.93–0.99, respectively. Notably, the predictive accuracy of all models improved as the number of input variables increased. To validate the applicability of these models within a Geographic Information Systems (GIS) environment, the best RFR-derived model was used to produce spatial prediction maps via the Inverse Distance Weighted (IDW) method. The generated maps closely aligned with the spatial distribution maps of the measured values. This study recommends further exploration and application of machine learning models, particularly Random Forest Regression, for predicting soil properties that are challenging to measure directly.

Keywords: Pedotransfer Functions, AL wadi Al-Hai, Random Forest Regression, Libyan soils.

*Corresponding Author: Younes Daw Z. Ezlit - Dep. of soil and water - Fac. Of Agric. - Univ. of Tripoli. - Libya.

Phone: +218917830015

E-mail: y.ezlit@uot.edu.ly

Received: 22/7/2024

Accepted: 15 /1/ 2025