

دراسة وتقدير قيم المادة العضوية في منطقة مختارة من البادية السورية باستخدام البيانات الاستشعارية والحقلية

عمر عبد الرزاق⁽¹⁾ - هاني محمد إبراهيم⁽²⁾ - وفاء العسكر الهفل⁽³⁾

⁽¹⁾ أستاذ في قسم الأراضي - كلية الهندسة الزراعية - جامعة الفرات.

⁽²⁾ دكتور في الهيئة العامة للاستشعار عن بعد - دمشق.

⁽³⁾ طالبة دكتوراه في قسم الأراضي - كلية الهندسة الزراعية - جامعة الفرات

الملخص:

أن الفهم الدقيق للتوزع المكاني والتغيرات الحاصلة في كمية المادة العضوية أمر ضروري لتخطيط استعمالات الأراضي والأنشطة الأخرى المتعلقة بإدارة الزراعة وحماية البيئة. تم في هذه الدراسة استخدام صور LANDSAT 8 وLANDSAT 5 لاشتقاق مجموعة من الدلائل الطيفية، إضافة لاستخدام نموذج الارتفاع الرقمي لاشتقاق مجموعة من الدلائل الطبوغرافية ذات الصلة بكمية المادة العضوية في التربة، كما تم استخدام مجموعة كبيرة من البيانات الحقلية الخاصة بقيم المادة العضوية في التربة المتزامنة مع تواريخ الصور الفضائية، والعمل على خوارزمية الغابات العشوائية Random Forest Algorithm RFA وهي إحدى خوارزميات تعلم الآلة لتحديد التوزع المكاني لقيم المادة العضوية في تربة منطقة الدراسة. تم دراسة التغيرات الحاصلة على كمية المادة العضوية في تربة منطقة الدراسة منذ العام 2003 وحتى 2022 بفارق زمني ثلاثة سنوات، حيث بينت الدراسة فعالية استخدام طريقة RFR، والبيانات الاستشعارية المختلفة في التنبؤ بكمية المادة العضوية على فترات زمنية مختلفة لدراسة التغيرات الحاصلة في المادة العضوية في تربة المنطقة المدروسة.

كلمات مفتاحية: المادة العضوية، الصور الفضائية، الغابات العشوائية، البادية

السورية، الاستشعار عن بعد

المقدمة والدراسة المرجعية:

تُعد التربة جزء أساسي من النظم البيئية الجافة، كم أنها تلعب دوراً حيوياً في دورة التبادل للمادة والطاقة بين الغلاف الجوي والمحيط الحيوي (Sparling وزملائه، 2006; Zhang وزملائه، 2018). تعتبر المادة العضوية Soil Organic Matter SOM مكون مهم من مكونات التربة حيث تؤثر على الخصوبة وتخزين الكربون فيها (Manlay وزملائه، 2007; Nocita وزملائه، 2013)، كما وتلعب دوراً هاماً في تحسين غلة المحاصيل الحقلية (Feng وزملائه، 2016; Mishra وزملائه، 2012) والحفاظ على استدامة الزراعة والأمن الغذائي (Were وزملائه، 2015)، لذلك فإن التنبؤ الدقيق وتحديد التوزيع المكاني للمادة العضوية في التربة أمر مهم لتعزيز تحسين خصوبة التربة وتطوير نظام بيئي إقليمي مستدام ولتنفيذ السياسات المختلفة لإدارة التربة (Meersmans وزملائه، 2008). على الرغم من أن العديد من الدراسات قد أجريت حول بناء نماذج تنبؤ بالمادة العضوية ورسم خرائط لتوزيعها المكاني إلا أنه يظل أنجاز خريطة دقيقة لمحتوى التربة من المادة العضوية تحدياً كبيراً نتيجة لصعوبات اختيار مصادر البيانات الاستشعارية المناسبة ونماذج التنبؤ الخاصة بمنطقة محددة. توجد العديد من الطرق التي طورت لرسم خرائط التوزيع المكاني للمادة العضوية في التربة وعلى سبيل المثال فإن الطرق التقليدية تطبق إجراءات المعايرة (Walkley وزملائه، 1934) لعينات التربة في جهاز Spectrophotometer (Van Raij وزملائه، 2001) لتحديد محتوى التربة من المادة العضوية ومع ذلك فإنه من الصعب الحصول على التوزيع المكاني لقيم المادة العضوية على مقاييس مكانية كبيرة بصورة كمية (Lagacherie وزملائه، 2008) وذلك نتيجة للحجم الكبير من الوقت والمال اللازمين لتنفيذ المسوحات الحقلية، بالإضافة إلى ذلك فإن القيم المقاسة للمادة العضوية في التربة تتأثر بصورة كبيرة بالخبرة الحقلية التي يجب أن يتمتع بها مساحو التربة (Bie وزملائه، 1971; Zhao وزملائه، 2013). تم استخدام مجموعة من الطرق الإحصائية الجغرافية Geostatistical (Matheron وزملائه، 1963) بصورة واسعة (Kriging and Cokriging) للتنبؤ بقيم المادة العضوية في التربة (Meersmans وزملائه، 2008;

Zhang وزملائه، 2012; Dai وزملائه، 2014) وعلى أي حال فإن تلك الطرق تتطلب جمع عينات تربة كثيفة بصورة فعالة (Dai وزملائه، 2014; Zhao وزملائه، 2010) وممثلة للمنطقة المدروسة بصورة جيدة (Ward وزملائه، 2019). أن التطور السريع في تكنولوجيا الاستشعار عن بعد ساعد في إيجاد طرق مختلفة ودقيقة للتنبؤ بقيم المادة العضوية في التربة باستخدام بيانات الاستشعار عن بعد المشتقة من صور التوابع الصناعية المختلفة، حيث اعتمدت أولى المحاولات للتنبؤ بالمادة العضوية في التربة باستخدام البيانات الاستشعارية على تحليل الانحدار الخطي المتعدد (Meersmans وزملائه، 2008; Takata وزملائه، 2007; Cheng وزملائه، 2004)، بينما يتم حالياً تطوير العديد من النماذج الرياضية والإحصائية التجريبية والتي تعطي نتائج تنبؤ دقيقة (Bogunovic وزملائه، 2018; Rosero-Vlasova وزملائه، 2018) وهذه النماذج عادة ما تربط بين الطرق الإحصائية المختلفة والمتغيرات التنبؤية المرتبطة بمحتوى التربة من المادة العضوية (Rosero-Vlasova وزملائه، 2018; Bobrovsky وزملائه، 2010) ومثال ذلك عندما يتم الدمج بين البيانات الاستشعارية والبيانات الطبوغرافية مع الأخذ بعين الاعتبار وجود مشكلات عديدة عند استخدام طرق Kriging وطرق الانحدار الخطي المتعدد ومنها ما هو متعلق بمشكلات من نوع Nonlinearity و Multicollinearity (Alvarez وزملائه، 2011; Zhao وزملائه، 2020)، وللتغلب على تلك القيود فإن تطبيق خوارزميات تعلم الآلة Machine Learning Algorithms مثل خوارزميات الغابات العشوائية Random Forest RF (Chen وزملائه، 2018; Qi وزملائه، 2017) والشبكات العصبية الاصطناعية Artificial Neural Network ANN (Dai وزملائه، 2014; Zhao وزملائه، 2010; Guo وزملائه، 2013; Zhao وزملائه، 2020; Fernandes وزملائه، 2019; R Core Team، 2013) وانحدار دعم المتجه Support Vector Regression SVR (Were وزملائه، 2015; Deiss وزملائه، 2020; Ballabio وزملائه، 2009) قد تطورت بصورة تدريجية في أبحاث علوم التربة بفضل الفوائد الهامة لتلك الخوارزميات مقارنة بالنهج السابق (تحسين دقة النموذج، زيادة كفاءة

التنفيذ، تبسيط الارتباط (Simplified Fitting). حسب كل من (Chen وزملائه، 2019; Kumar وزملائه، 2011) فإنه لا يوجد نموذج رياضي وحيد قابل للتطبيق عالمياً للتنبؤ بقيم المادة العضوية في التربة في مناطق مختلفة وذلك بفعل التغيرات المكانية في الظروف المناخية ومواد التربة الأم والنشاطات البشرية وإدارة استعمالات الأراضي. أشارت العديد من الدراسات أن المادة العضوية في التربة حساسة عبر منطقة الطيف الكهرومغناطيسي المرئي وتحت الأحمر وتحت الأحمر قصير الموجة (400 - 2500 نانوميتر) (Liu وزملائه، 2008; Chen وزملائه، 2000) حيث أن محتوى التربة من المادة العضوية يمكن أن يحدد من خلال الاختلافات في خصائص الانعكاس الطيفي للتربة (Mccarty وزملائه، 2002; Bilgili وزملائه، 2010). قام كل من (Zhang وزملائه، 2021) بدراسة لتقييم عمل النماذج التالية RF و SVR و ANN في التنبؤ بقيم المادة العضوية ورسم خرائط لها في سهول Songnen في الصين، حيث بينت النتائج أن النموذج RF أعطى أقل قيمة خطأ (Root Mean Squared Error RMSE = 0.68 %) وأعلى قيمة معامل تحديد ($R^2 = 0.67$) من النموذجين ANN و RF كما وضحت الدراسة أن استخدام صور التابع الصناعي Sentinel-2A أعطت نتائج أفضل من استخدام بيانات التابع الصناعي MODIS بفضل قدرتي التمييز المكاني والطيفي العالية التي تتمتع بها صور التابع الصناعي Sentinel-2A مقارنة بصور التابع الصناعي MODIS. بين كل من (Farooq وزملائه، 2022) في دراسة للتنبؤ بمخزون التربة من الكربون العضوي في منطقة الهيمالايا في الهند ورسم خرائط توزع مكاني لها أن خوارزمية RF أعطت نتائج أفضل ($RMSE = 8.21$) من الطرق الإحصائية الجغرافية Geostatistical كطريقة Ordinary Kriging OK ($RMSE = 15.6$) وطريقة Regression Kriging RK ($RMSE = 17.73$)، في التنبؤ بمخزون التربة من المادة العضوية ورسم خرائط التوزيع المكاني لها في منطقة الدراسة، علماً أنه تم في هذه الدراسة استخدام مجموعة من البيانات الحقلية عن مخزون التربة من الكربون العضوي، إضافة لعدد من المتغيرات المستقلة (دلائل طيفية مشتقة من البيانات الاستشعارية، دلائل طبوغرافية) الخاصة بالتنبؤ بمخزون الكربون العضوي،

وعلاوة على ذلك فقد أكد كل من (Pham وزملائه، 2019; Wiesmeier وزملائه، 2011) على أن خوارزميات التعلم الآلي Regression Tree و SVR و Cubist و RF من بين أمور أخرى تم تطويرها في السنوات الماضية تقوم على ربط العوامل البيئية المختلفة مع خصائص التربة المتعددة لالتقاط العلاقات غير الخطية القائمة بين هذه المتغيرات مما يؤدي إلى إعطاء تنبؤ أفضل ونتائج دقيقة.

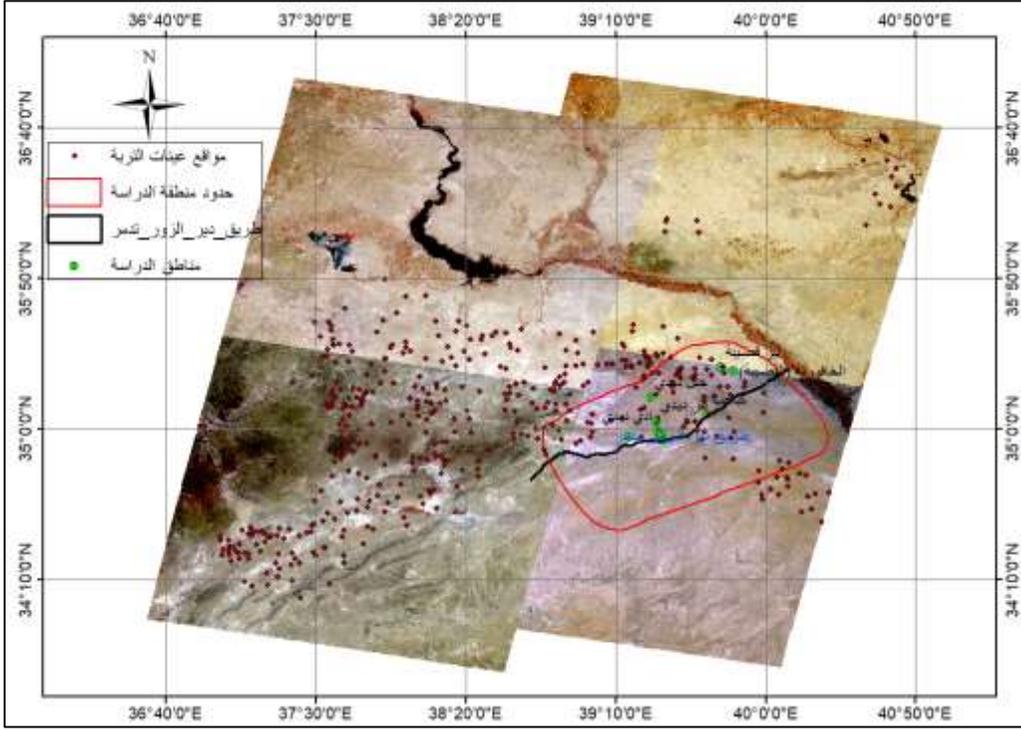
أهداف الدراسة:

- 1- تقدير قيم المادة العضوية في ترب منطقة الدراسة باستخدام البيانات الاستشعارية والحقلية.
- 2- دراسة تغيرات متوسطات قيم المادة العضوية في ترب منطقة الدراسة في أربعة صفوف من الارتفاع عن سطح البحر للأعوام من 2003 وحتى 2022 بفواصل زمني بين كل مرحلة وأخرى ثلاثة سنوات.

مواد وطرق البحث:

1- منطقة الدراسة

تُعد منطقة الدراسة جزءاً من البادية السورية وتتوزع بين محافظتي دير الزور وحمص، حيث أنها تمتد بين خطي طول (38.762535 و 40.350040 درجة) وخطي عرض (34.433201 و 35.499757 درجة)، وتبلغ مساحة منطقة الدراسة حوالي 10371.6 كم²، حيث يحدها من الشمال الشرقي وادي الفرات، الشكل (1) يبين حدود منطقة الدراسة في البادية السورية، إضافة إلى التوزيع المكاني لنقاط عينات التربة الحقلية اللازمة لغرض النمذجة الرياضية في الدراسة فيما بعد.



الشكل (1) يبين حدود منطقة الدراسة وتوزع نقاط التربة الحقلية ضمنها

2- المواد والطرائق

2-1-البيانات الحقلية

شملت البيانات الحقلية استخدام 550 عينة ترابية حقلية سطحية، تم جمعها في منطقة البادية السورية خلال مشروع مسح البادية بالتعاون مع اكساد، حيث نفذت الأعمال الحقلية بين عامي 2003 و 2004، وتم توثيق الإحداثيات لكل نقطة باستخدام جهاز (The Global Positioning System) GPS، بالإضافة إلى 14 مقطع تربة تم تنفيذها في محافظتي الحسكة و الرقة حيث تراوحت المسافة بين نقاط عينات التربة الحقلية بين 1,5 إلى 2 كم، وتوزعت مكانياً بصورة متجانسة على كامل مساحة البادية ضمن منطقة الدراسة وخارج منطقة الدراسة، حيث تم استخدام بيانات النقاط الحقلية خارج منطقة الدراسة لزيادة حجم العينة الحقلية والتي سوف تعطي نماذج ومعادلات انحدار خطية متعددة أكثر مصداقية وتمثيل لأرض الواقع وقابلة للتطبيق في المساحات

ذات الظروف المتشابهة من حيث طبيعة المناخ والتربة وشكل التضاريس، إضافة للتشابه الحاصل بين منطقة الدراسة والمناطق المحيطة من حيث طبيعة المناخ و التربة واشكال استعمالات الأراضي والغطاء الأرضي المتواجدة.

2-2- البيانات الاستشعارية

استخدمت في هذه الدراسة مجموعة من البيانات الاستشعارية مختلفة الخصائص وهي على التوالي:

2-2-1- صور فضائية للتابعين الصناعيين LANDSAT 5 وLANDSAT 8 بقدرة تمييز مكاني (30*30م) في المجالات الطيفية التالية (الأزرق، الأخضر، الأحمر، وتحت الأحمر القريب، تحت الأحمر القريب قصير الموجة) غطت كامل منطقة الدراسة خلال مراحل العمل، تم الحصول عليها من الموقع الإلكتروني <http://earthexplorer.usgs.gov>. تتكون صور التابع الصناعي LANDSAT 8 من إحدى عشر قناة طيفية منها قناتين طيفيتين في المجال تحت الأحمر الحراري، وتتميز بقدرة تمييز زمانية كل 16 يوم، بينما تتكون صور LANDSAT 5 من سبعة قنوات طيفية منها قناة طيفية في المجال تحت الأحمر الحراري. تم في هذه الدراسة جمع القنوات الطيفية (b2, b3, b4, b5, b6, b7) و (b1, b2, b3, b4, b5,) التي تعبر عن المجالات الطيفية (الأزرق، الأخضر، الأحمر، وتحت الأحمر القريب، تحت الأحمر قصير الموجة) في صور LANDSAT 8 و صور LANDSAT 5 المستخدمة على التوالي في صورة واحدة Layer Stack، ومن ثم تحويل قيم الصور الرقمية من Digital Number (DN) إلى قيم الانعكاس الطيفي Spectral Radiance ومن ثم إلى قيم الانعكاسية الطيفية Spectral Reflectance في كل قناة باستخدام المعادلة التالية للحصول على قيم ينحصر مجالها بين (0 و 1) لاستخدامها في النماذج الرياضية المختلفة.

$$P_{\lambda} = \frac{\pi * L_{\lambda}}{ESUN_{\lambda} * \text{Cos}\theta * d_r}$$

حيث أن (L_{λ}) الانعكاس الطيفي لكل قناة طيفية، $(ESUN_{\lambda})$ متوسط الأشعة الشمسية الواردة خارج الغلاف الجوي لكل قناة طيفية وواحدتها (واط/م²/ميكرومتر)، $(\cos\theta)$ تجيب زاوية ورود الأشعة الشمسية، (d_r) مقلوب مربع المسافة النسبية بين الشمس والأرض.

2-2-2- نموذج الارتفاع الرقمي (DEM) Digital Elevation Model الذي يُعد تمثيل لتضاريس سطح الأرض، كما يعتبر من أكثر أنواع البيانات الجغرافية أهمية لاستخدامه في مجالات ودراسات مختلفة. تم في هذه الدراسة الحصول على بيانات نموذج الارتفاع الرقمي لمنطقة الدراسة بقدرة تمييز مكاني 30م من خلال الموقع الإلكتروني المجاني <http://earthexplorer.usgs.gov>.

2-3-2- تم حساب عدد من الدلائل الطيفية باستخدام الصور الفضائية **LANDSAT 5 و 8** الملتقطة بتواريخ مختلفة والتي غطت كامل منطقة الدراسة حتى تتوافق ومواعيد تنفيذ الاعمال الحقلية واخذ عينات التربة، حيث تم حساب قيم كل دليل طيفي للصور الفضائية الملتقطة خلال شهر نيسان وحتى شهر أيلول لكل من الأعوام 2003 - 2004 - 2013 والتي تغطي نقاط بيانات التربة الحقلية، ومن ثم عمل متوسط لقيم هذا الدليل الطيفي للأشهر الستة الممتدة من شهر نيسان إلى شهر أيلول، للحصول على متوسط قيم الدليل الطيفي المحسوب والذي سيتم أدخله في النمذجة الرياضية فيما بعد وتعطى هذه الدلائل الطيفية كما يلي:

2-3-1- دليل التغيرات النباتية المعدل **Normalized Differences Vegetation Index (NDVI)**

ويمثل الفرق في الانعكاس بين القناتين الطيفيتين المرئية الحمراء وتحت الحمراء مقسوما على مجموع القناتين المرئية الحمراء وتحت الحمراء، وتتراوح قيم الدليل النباتي NDVI بين 1- إلى 1+ حيث تمتد القيم الفعلية لهذا الدليل من 0.1 للأراضي الخالية من الغطاء النباتي (تربة) إلى 0.9 للغطاء النباتي جيد النمو وتعطى معادلة الدليل النباتي كما يلي:

$$NDVI = \frac{NIR - R}{NIR + R}$$

حيث أن NIR المجال الطيفي تحت الأحمر القريب، R المجال الطيفي الأحمر.

2-3-2- دليل درجة حرارة سطح الأرض Land Surface Temperature LST

تم استخدام دليل LST لما له من تأثير كبير على عملية المعدنة للمادة العضوية الحاصلة في التربة من جهة، ولتأثير درجة حرارة سطح الأرض على نمو النباتات الحولية والموسمية من جهة أخرى. تم حساب درجة حرارة سطح الأرض LST باستخدام معادلة Plank كما يلي:

$$LST = \frac{K_2}{LN \left[\left(\frac{\epsilon_{NB} K_1}{L_\lambda} \right) + 1 \right]}$$

حيث أن K_1 و K_2 ثابتت معايرة خاصة بصور LANDSAT وتعطى من خلال الملف الملحق بالصورة الفضائية Header File، L_λ الانبعاثية الطيفية للقناة الطيفية الحرارية Thermal Band في صور LANDSAT بوحدة (واط/م² /ستراديان/ميكروميتر)، ϵ_{NB} انبعاثية سطح الأرض Land Surface Emissivity ويتم حساب معاملات المعادلة السابقة كما يلي:

- L_λ الانبعاثية الطيفية للقناة الطيفية الحرارية Thermal Band في صور LANDSAT بوحدة (واط/م² /ستراديان/ميكروميتر):

$$L_\lambda = \left(\frac{L_{MAX} - L_{MIN}}{255} \right) * DN + L_{MIN}$$

حيث أن L_{MAX} و L_{MIN} ثابتت معايرة خاصة بكل قناة طيفية من قنوات الصور الفضائية LANDSAT وتعطى من خلال الملف الملحق بالصورة الفضائية Header File، DN Digital Number وهي القيمة الرقمية لكل بكسل من بكسلات الصور الفضائية المستخدمة في الدراسة من نوع LANDSAT.

- انبعاثية سطح الأرض Land Surface Emissivity ϵ_{NB} :

تعرف الانبعاثية السطحية بأنها نسبة الأشعة الفعلية المنبعثة عن سطح ما إلى الأشعة المنبعثة عن الجسم الأسود عند درجة حرارة سطح محددة (نفس درجة الحرارة).

حيث تم استخدام عتبة القرينة النباتية NDVI Thresholds Method لتقدير قيم انبعاثية سطح الأرض من صور LANDSAT كما في المعادلات التالية:

$$\varepsilon_i = \begin{cases} a_i * p_{red} + b_i & NDVI < 0.2 \\ \varepsilon_{v,i} * P_v + \varepsilon_{s,i} * (1 - P_v) + C_i & 0.2 \leq NDVI \leq 0.5 \\ \varepsilon_{v,i} + C_i & NDVI > 0.5 \end{cases}$$

تشتق القرينة النباتية P_v Vegetation Fraction من القرين النباتية NDVI والتي تشتق بدورها في صور LANDSAT من القناتين الطيفيتين الأحمر وتحت الأحمر.

$$P_v = \left[\frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}} \right]^2$$

حيث $NDVI_{min}$ القيمة الصغرى للقرينة النباتية NDVI في الصورة الفضائية و $NDVI_{max}$ القيمة الكبرى للقرينة النباتية NDVI في الصورة الفضائية.

C_i : مصطلح يأخذ بعين الاعتبار تأثير التضاريس بفعل خشونة السطح surface roughness (في السطوح المستوية فإن $C_i = 0$)، حيث يتم تقدير قيمة C_i كما في المعادلة التالية:

$$C_i = (1 - \varepsilon_{s,i}) \times \varepsilon_{v,i} \times F' \times (1 - P_v)$$

حيث أن F' عامل هندسي Geometric Factor يتراوح بين 0 و 1 ويعتمد على التوزيع الهندسي للسطح ويعطى بصورة أنموذجية $F' = 0.55$.

تم حساب انبعاثية الغطاء النباتي ε_v Vegetation Emissivity وانبعاثية التربة Soil Emissivity ε_s من MODIS UCSB (University of California, Santa Barbara) emissivity library، الجدول التالي يبين قيم انبعاثية الغطاء النباتي وانبعاثية التربة

القنوات الطيفية	ε_s التربة	ε_v الغطاء النباتي
Thermal Band	0.9668	0.9863

يُعد بكسل الصورة LANDSAT بكسل تربة جرداء (bare soil) ($P_v = 0$) عندما يكون $NDVI < 0.2$ وفي هذه الحالة يتم تقدير قيم انبعاثية سطح الأرض من خلال علاقة

تجريبية خطية باستخدام قيم الانعكاسية الطيفية للقناة الحمراء p_{red} ، وتعطى قيم انبعاشية سطح التربة الخالية من الغطاء النباتي للقناة الحرارية في صورة LANDSAT كما يلي:

$$\varepsilon_i = 0.973 - 0.047 \times p_{red}$$

حيث: P_{red} انعكاس الأشعة عن سطح الأرض في المجال الأحمر.

2-3-3- دلائل الألبيدو السطحي Surface Albedo

يعرف الألبيدو السطحي بأنه نسبة الأشعة الشمسية المنعكسة إلى الأشعة الشمسية قصيرة الموجة الواردة على المساحة شبه الكروية للأرض، ويتم حساب قيم الألبيدو السطحي في صور LANDSAT كما يلي:

- حساب الألبيدو السطحي α Surface Albedo

$$\alpha = \frac{\alpha_{toa} - \alpha_{path\ radiance}}{(\tau_{sw})^2}$$

$\alpha_{path\ radiance}$ متوسط الحصة من الإشعاع الشمسي الوارد عبر كل القنوات الطيفية والتي تتبعثر عائدة إلى القمر الصناعي قبل أن تصل إلى سطح الأرض، τ_{sw} نفاذية الغلاف الجوي، α_{toa} Top of Atmosphere Albedo قيم الألبيدو خارج الغلاف الجوي، ويتم حساب مكونات المعادلة السابقة كما يلي:

$$\alpha_{toa} = \sum (\omega_\lambda * \rho_\lambda)$$

ρ_λ انعكاسية القنوات الطيفية Reflectivity، ω_λ معامل تثقيل Weighting Coefficient لكل قناة طيفية انعكاسية من القنوات الطيفية الانعكاسية في صور LANDSAT

2-4- تم حساب عدد من الدلائل الطبوغرافية باستخدام نموذج الارتفاع الرقمي Digital Elevation Model DEM والتي تعطى كمايلي:

2-4-1- نموذج الارتفاع الرقمي (DEM) Digital Elevation Model الذي يُعد تمثيل لتضاريس سطح الأرض المستمرة، كما يعتبر من أكثر أنواع البيانات الجغرافية أهمية لاستخدامه في مجالات ودراسات مختلفة.

2-4-2- الانحدار Slope والذي يشير إلى زاوية ميل سطح الأرض عن الأفق.

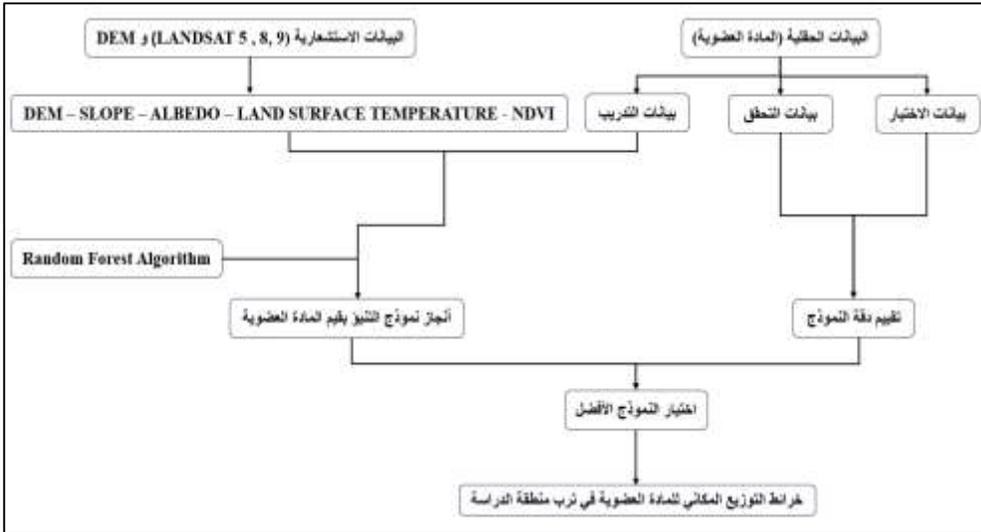
2-5- تم تصدير متوسط قيم كل من الدلائل الطيفية (ALBEDO، LAND SURFACE TEMPERATURE، NDVI) المحسوبة خلال الستة أشهر من كل عام (2003 - 2004، و 2013) والتي شملت شهر نيسان وحتى شهر أيلول، إضافة لتصدير قيم الدلائل الطبوغرافية (% Slope، DEM) إلى ما يقابلها من قيم المادة العضوية المقاسة حقلياً في 564 نقطة حقلية مأخوذة في الأعوام (2003 - 2004) و (2013) والموثقة الاحداثيات الجغرافية باستخدام برنامج نظم المعلومات الجغرافية ArcGIS للحصول فيما بعد على جدول أكسل لبيانات 564 نقطة حقلية موزعة على كامل منطقة الدراسة وخارج منطقة الدراسة وتضم قيم المادة العضوية الحقلية وما يقابلها من قيم كل من الدلائل الطيفية وقيم الدلائل الطبوغرافية المحسوبة من الصور الفضائية LANDSAT 5 و LANDSAT 8 ليصار فيما بعد لاستخدامها في تقدير قيم المادة العضوية في ترب منطقة الدراسة خلال الفترات الزمنية المتتالية.

2-6- بناء النموذج الرياضي الخاص بالتنبؤ بقيم المادة العضوية في ترب منطقة الدراسة

تم في هذه الدراسة استخدام خوارزمية الغابات العشوائية Random Forest Regressor وهي إحدى خوارزميات تعليم الآلة المراقب Supervised Machine Learning، وكل من البيانات الطيفية والطبوغرافية (SLOPE، DEM، NDVI، LAND SURFACE TEMPERATURE، ALBEDO) المشتقة من الصور الفضائية LANDSAT، إضافة للبيانات الحقلية الخاصة بقيم المادة العضوية، حيث تقوم خوارزمية RFR على انشاء مجموعة من أشجار القرار Decision Trees بصورة عشوائية من خلال تحديد جزء عشوائي من بيانات التدريب وبناء شجرة القرار الخاصة بهذه المجموعة من البيانات ومن ثم أخذ جزء آخر من البيانات بصورة عشوائية وبناء شجرة قرار أخرى وهكذا لباقي بيانات التدريب، بحيث أن كل شجرة من أشجار القرار تأخذ جزء من بيانات التدريب (أجزاء متساوية أو غير متساوية) وتتخذ في النهاية القرار المناسب الخاص ببيانات التدريب، ليتم فيما بعد تجميع القرارات من أشجار القرار المختلفة التي تم تشكيلها أثناء عمل خوارزمية RFR لتحديد التنبؤ النهائي في النموذج.

يتم في خوارزمية RFR استخدام مجموعة من العوامل تسمى Hyperparameters والتي لا يمكن تحديد قيمها من خلال التعلم من بيانات التدريب المستخدمة في النموذج ولكن يتم تقديرها مسبقاً قبل بدء عملية التعلم باستخدام خوارزمية Grid Search CV وهي طريقة برمجية تعمل على ضبط قيم Hyperparameters، حيث تعمل على اختيار قيم Hyperparameters المثالية من مجموعة من قيم Hyperparameters التي تزود بها خوارزمية RFR قبيل بدء عملها على بيانات التدريب ومن هذه العوامل التي يجب تحديدها العامل المسمى n-estimators والذي يعبر عن عدد أشجار القرار المستخدمة في الحساب، إضافة إلى العامل max-depth والذي يحدد مدى عمق كل شجرة من أشجار القرار العاملة في خوارزمية RFR. تم استخدام القيم التالية n-estimators = [30, 50, 100, 150, 200, 250, 300, 350, 400] و max-depth = [None, 5, 10, 15, 20, 25, 30, 35, 40] عمل خوارزمية RFR على بيانات التدريب الخاصة بقيم المادة العضوية في التربة، ولاختيار أفضل القيم اللازمة لعمل خوارزمية RFR وتطبيقها على كامل بيانات التدريب وللتنبؤ بالقيم المستقبلية فقد تم استخدام الطريقة الإحصائية Cross-Validation والتي تقوم في عملها على تقسيم بيانات التدريب إلى عدد من مجموعات البيانات المتساوية والتي يطلق عليها K Fold Cross-Validation وفي دراستنا تم تقسيم البيانات إلى خمسة مجموعات K = 5 بحيث أن خوارزمية RFR تعمل على بيانات التدريب لأربعة مجموعات Training Data ومجموعة البيانات الخامسة تستخدم كبيانات تحقق Testing Data ويتكرر عمل خوارزمية RFR بحيث كل مجموعة من مجموعات البيانات الخمسة تعمل مرة واحدة كبيانات تحقق وبقية المجموعات كبيانات تدريب، وبالتوازي مع عمل Cross-Validation فقد تم استخدام مقاييس الأداء التالية معامل التحديد Coefficient of Determination R² وهو مقياس إحصائي يمثل النسبة المئوية للتباين في المتغير التابع والذي يتم تفسيره من خلال التغيرات الحاصلة في المتغير المستقل أو مجموعة المتغيرات المستقلة، و جذر متوسط مربع الفروق بين القيم الحقيقية والمتوقعة Root Mean Squared Error RMSE لاختيار النموذج RFR

الأفضل وتطبيقه على كامل بيانات التدريب ليكون جاهز للتنبؤ بالقيم المستقبلية التي لم يتدرب عليها النموذج علماً أنه في بداية العمل وقبل اختيار النموذج الأفضل بناءً على Hyperparameters الأفضل تم تقسيم البيانات إلى مجموعتين من البيانات هما بيانات التدريب Training Data بنسبة 80 % من البيانات الكلية والتي تم من خلالها اختيار النموذج الأفضل للتنبؤ كما سبق شرحه و 20 % كبيانات اختبار Testing Data يتم استخدامها كبيانات لم يسبق ان تدرب عليها النموذج وذلك لاختبار دقة عمل النموذج على بيانات جديدة. الشكل (1) يوضح مخطط عمل النموذج الرياضي الخاص بتقدير المادة العضوية في ترب منطقة الدراسة.



الشكل (1) يبين مخطط عمل نموذج تقدير المادة العضوية

أن استخدام خوارزميات الغابات العشوائية RFR يتضمن مجموعة من المزايا والفوائد والتي قد لا تتواجد في كثير من خوارزميات تعلم الآلة المستخدمة في التصنيف أو التنبؤ ومنها على سبيل المثال: 1- السرعة في إنجاز عملية التنبؤ باستخدام بيانات التدريب ومن ثم التحقق من فعالية ودقة النموذج باستخدام بيانات التحقق، 2- تستطيع خوارزمية RFR وبصورة فعالة من تجنب حدوث ما يسمى بالأفراط بالتعلم Overfitting بفضل توظيف استراتيجية العينة العشوائية وبالتالي تجنب عدم حدوث استقرار في النموذج عند اجراء عمليات التحقق من دقة عمل النموذج باستخدام بيانات التحقق

Testing Data (Grimm وزملائه، 2008; Meng وزملائه، 2020)، 3- تعمل خوارزمية RFR على تجنب مشاكل العلاقات الخطية المتعددة Multicollinearity بين العوامل المستقلة Features الداخلة في التنبؤ بصورة فعالة، كما وتدعم العوامل المستقلة عالية الابعاد (Ye وزملائه، 2013; Menze وزملائه، 2009)، 4- تتمتع خوارزمية RFR بقدرة كبيرة على التعامل مع القيم المتطرفة ضمن بيانات المتغيرات الداخلة إلى الخوارزمية، حيث يتعامل مع القيم المتطرفة من خلال استخدام طريقة الربط (Kuter، 2021)، 5- تستطيع خوارزمية RFR تبيان أهمية كل متغير من المتغيرات التفسيرية الداخلة إلى الخوارزمية وتحديد مدى مساهمتها في النموذج المنجز وفي عوامل التنبؤ الرئيسية بقيم العامل التابع (Kuter، 2021; Zhang وزملائه، 2012)، 6- تساعد العلاقات الرياضية الهرمية التي تتمتع بها خوارزمية RFR في إظهار استقرار جيد في عمل الخوارزمية عند التعامل مع البيانات المفقودة أو غير المتوازنة (Heung وزملائه، 2014). نظراً للتباين المكاني في الظروف المناخية، مكونات التربة الأصلية، الأنشطة البشرية، وإدارة استعمالات الأراضي فقد تختلف دقة التنبؤ بين النماذج الرياضية والاحصائية المختلفة من منطقة دراسية إلى منطقة دراسية أخرى وبالتالي لا توجد خوارزمية واحدة قابلة للتطبيق عالمياً لمناطق دراسية مختلفة (Chen وزملائه، 2019; Kumar وزملائه، 2011).

النتائج والمناقشة

تبين الاحصائيات الوصفية لمحتوى التربة من المادة العضوية المقاسة حقلياً والتي تم تنفيذها باستخدام برنامج Python ما يلي: تراوحت قيم المادة العضوية المقاسة حقلياً في كامل البيانات المدخلة إلى النموذج بين 0 و 4.7 %، كما بلغ المتوسط 0.9924 %، في حين بلغ الانحراف المعياري لقيم المادة العضوية في التربة 0.64 % . بلغ عدد القيم من المادة العضوية المقاسة حقلياً المستخدمة في تدريب النموذج Training 447Points قيمة، حيث تراوحت قيم المادة العضوية في بيانات التدريب بين 0.067 و 4.7 %، والقيمة المتوسطة 0.9976 %، في حين بلغ الانحراف المعياري 0.6283

% . بلغ عدد القيم من المادة العضوية المقاسة حقلياً المستخدمة في اختبار النموذج Testing Points 112 قيمة، حيث تراوحت قيم المادة العضوية في بيانات الاختبار بين 0 و 3.82 %، والقيمة المتوسطة 0.9719 %، في حين بلغ الانحراف المعياري 0.6785 % . الجدول التالي (1) يبين يوضح البيانات الوصفية لقيم المادة العضوية المقاسة حقلياً المستخدمة في نموذج RFR.

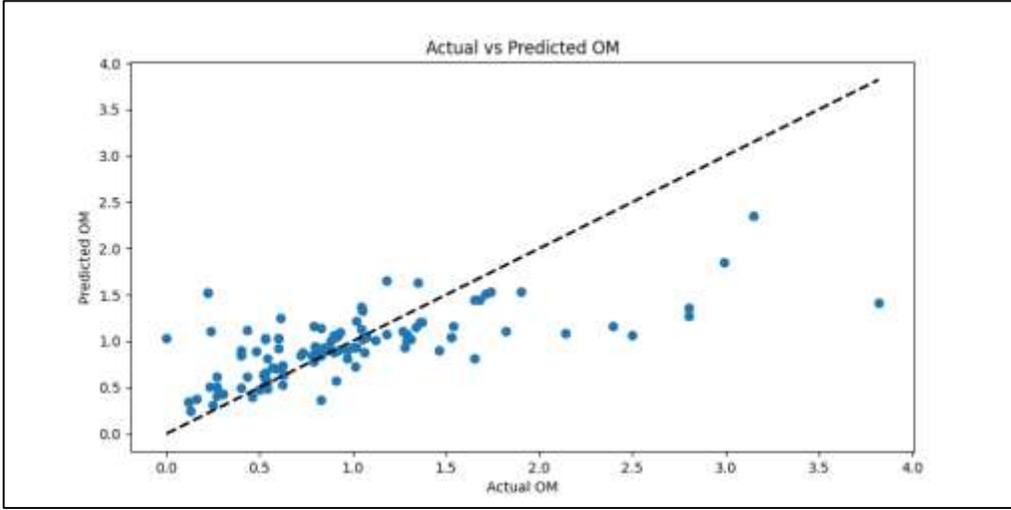
جدول (1) يبين البيانات الوصفية لقيم المادة العضوية الحقلية المستخدمة في النموذج

بيانات المادة العضوية	العدد	الحد الأدنى %	الحد الأعلى %	الانحراف المعياري %	المتوسط %
كامل البيانات	559	0	4.7	0.6381	0.9924
بيانات التدريب	447	0.067	4.7	0.6283	0.9976
بيانات الاختبار	112	0	3.82	0.6785	0.9719

اختيار قيم Hyperparameters باستخدام خوارزمية Grid Search CV لضبط العوامل الداخلة في عمل نموذج RFR، حيث بينت الدراسة ومن خلال عمل خوارزمية Grid Search CV التي تُعد طريقة برمجية تعمل على ضبط قيم Hyperparameters قبل بدء عملية التعلم أن القيم المثالية لكل من العامل n-estimators والذي يعبر عن عدد أشجار القرار المستخدمة في الحساب، إضافة إلى العامل max-depth والذي يحدد مدى عمق كل شجرة من أشجار القرار العاملة في خوارزمية RFR كانت على التوالي كما يلي: Best parameters: {'max_depth': 15, 'n_estimators': 100}. تم بدايةً تقسيم البيانات الحقلية للمادة العضوية في ترب منطقة الدراسة باستخدام خوارزمية تقسيم البيانات بصورة عشوائية والمسمى Train_Test_Split في مكتبة Sklearn العاملة في لغة البرمجة Python إلى مجموعتين من البيانات بنسبة 80 % لبيانات التدريب Training Data و 20 % لبيانات الاختبار Testing Data. من خلال استخدام الطريقة الإحصائية Cross-

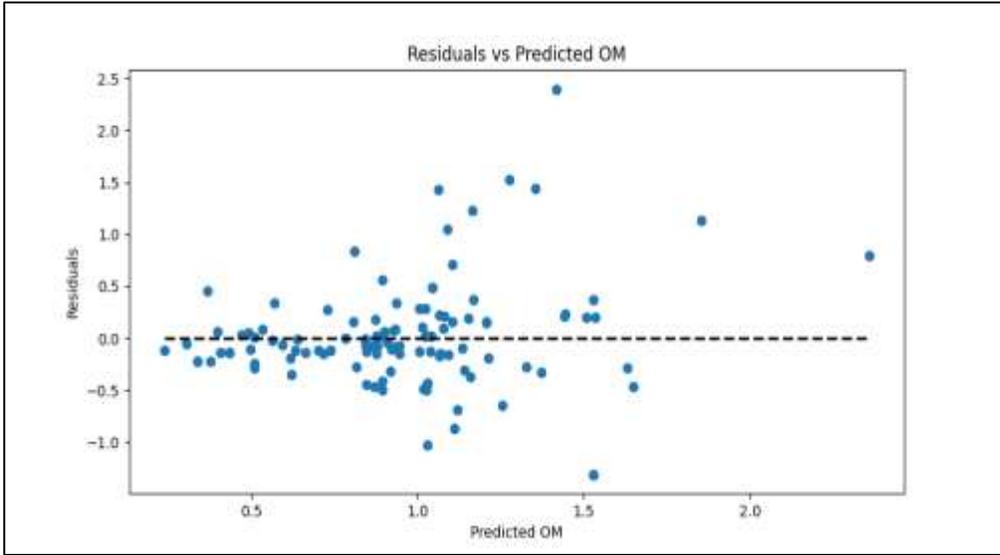
Validation والتي تم فيها تقسيم بيانات التدريب إلى خمسة مجموعات $K = 5$ بحيث عملت خوارزمية RFR في كل مرة على أربع مجموعات تدريب Training Points وتركت مجموعة البيانات الخامسة كبيانات اختبار Testing Data ، حيث تكررت العملية خمس مرات وفي كل مرة يتم استخدام مجموعة مختلفة من البيانات كبيانات اختبار وبقيّة المجموعات كبيانات تدريب، حيث تم في كل تكرار حساب معامل الارتباط بيرسون R ، و جذر متوسط مربع الفروق RMSE بين قيم المادة العضوية المتنبئ بها من خلال استخدام مجموعات التدريب الأربعة في النموذج RFR وقيم المادة العضوية المقاسة حقلياً والتي تمثل بيانات مجموعة الاختبار الخامسة، ومن ثم عمل متوسط لقيم كل من R^2 ، RMSE لجميع التكرارات، حيث بلغت القيم الوسطى $R = 0.406$ ، $RMSE = 0.57$.

بلغت قيمة معامل الارتباط بيرسون Pearson Correlation $R = 0.661$ بين قيم المادة العضوية المتنبئ بها Predicted OM وقيم المادة العضوية المقاسة حقلياً Actual OM في بيانات الاختبار Testing Points البالغ عددها 112 قيمة والتي لم تستخدم في عملية تدريب النموذج في خوارزمية RFR. كما بلغت قيمة جذر متوسط مربع الفروق بين القيم الحقيقية والمتوقعة للمادة العضوية في التربة $RMSE = 0.516$. الشكل التالي (2) يبين التوزع البياني لقيم المادة العضوية المتنبئ بها من خلال تطبيق خوارزمية RFR وقيم المادة العضوية المقاسة حقلياً في بيانات الاختبار.

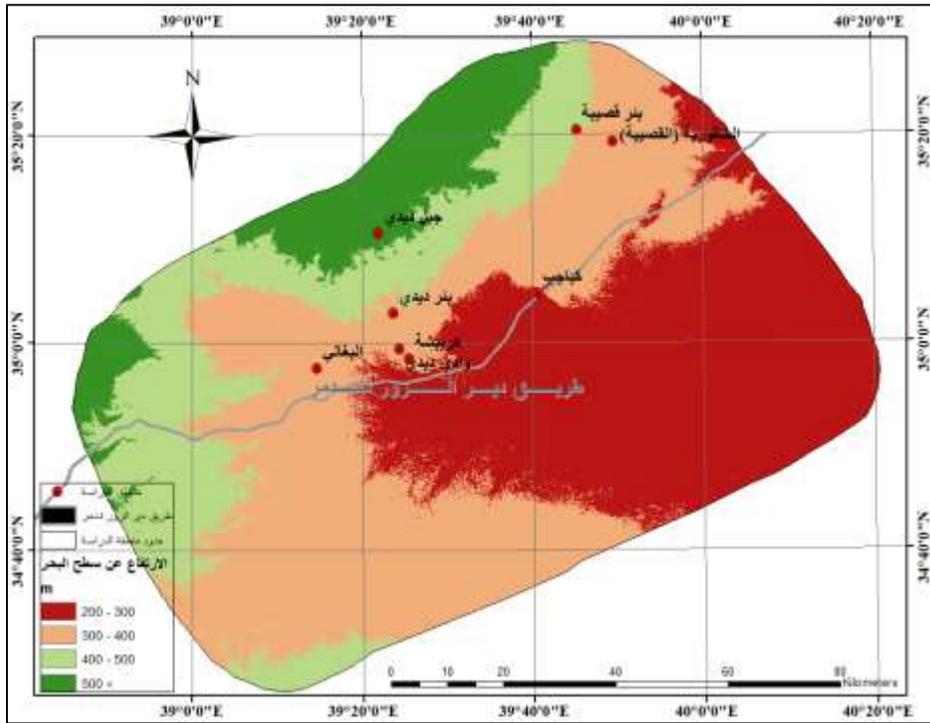


الشكل (2) يبين التوزيع البياني للقيم المتنبئ بها من المادة العضوية والقيم المقاسة حقلياً

كما يبين الشكل (3) التالي التوزيع البياني لقيم المادة العضوية المتنبئ بها من خلال تطبيق خوارزمية RFR من جهة والبواقي Residuals (الفرق بين القيم المتنبئ بها والقيم المقاسة حقلياً) من جهة أخرى وهو مخطط بياني يساعد في تحديد وجود أي أنماط معينة في القيم المتبقية والتي قد تشير إلى أن تنبؤات النموذج مرتفعة أو منخفضة بصورة منهجية بالنسبة لقيم معينة من المادة العضوية.



الشكل (3) يبين التوزيع البياني للقيم المتنبئ بها من المادة العضوية وقيم البواقي شكلت مساحة قيم الارتفاع عن سطح البحر بين (200 – 300م) ما نسبته 46.5 % من مساحة منطقة الدراسة، ومساحة قيم الارتفاع (300 – 400م) ما نسبته 52.6 %، ومساحة قيم الارتفاع (400 – 500م) ما نسبته 14.4 %، في حين بلغت النسبة المئوية لقيم الارتفاع عن سطح البحر (< 500م) ما نسبته 0.03 % من مساحة منطقة الدراسة. الشكل (4) يبين التوزيع المكاني لصفوف قيم الارتفاع عن سطح البحر في منطقة الدراسة.



الشكل (4) يبين التوزيع المكاني لصفوف الارتفاع عن سطح البحر

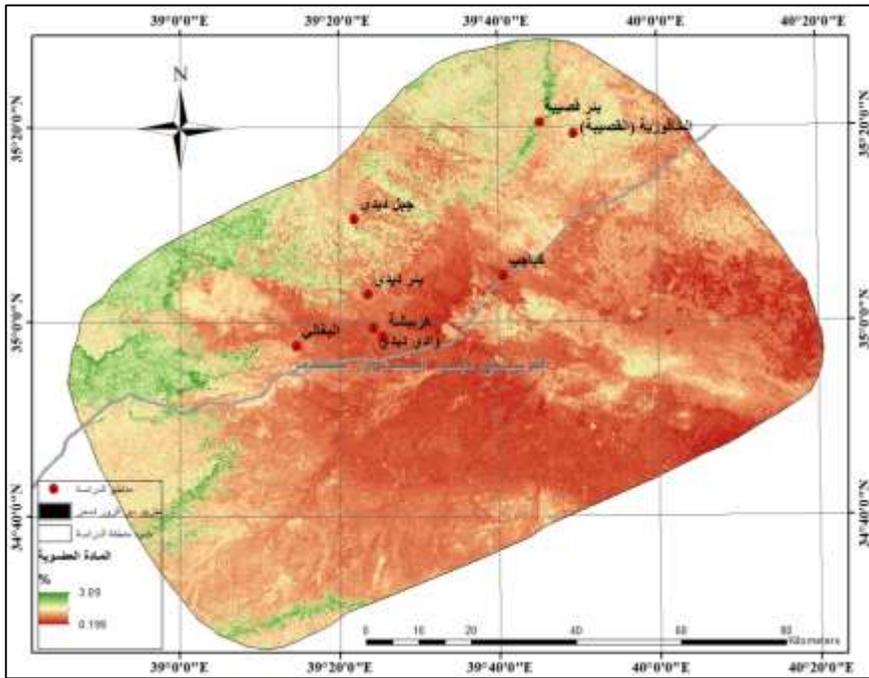
يلاحظ من الشكل (4) أن المناطق المرتفعة نسبياً بين (400 – 500م) و (< 500م) تنتشر في الأجزاء الغربية والجنوبية الغربية من منطقة الدراسة، بينما تتواجد المناطق متوسطة الارتفاع نسبياً (300 – 400م) في الأجزاء الوسطى، والجنوبية الشرقية من منطقة الدراسة، في حين تسود المناطق الأقل ارتفاعاً (200 – 300م) في الأجزاء الشمالية الشرقية وبمحاذاة وادي الفرات من الجهة الشمالية لمنطقة الدراسة. يلاحظ من الشكل (4) أن كل من منطقة هريشة والبالغية ومنطقة الخافورية، إضافة لمنطقة بئر ديدي ترتفع عن سطح البحر بين (300 – 400م)، بينما تقع منطقة كباجب ضمن نطاق الارتفاع عن سطح البحر بين (200 – 300م). يبين الشكل (5) متوسطات قيم المادة العضوية للارتفاعات بين (200 – 300م)، (300 – 400م)، (400 – 500م)، (< 500م) في منطقة الدراسة للأعوام من 2003- 2004 وحتى 2022.



الشكل (5) يبين التوزيع البياني لقيم المادة العضوية للارتفاعات المختلفة

يلاحظ من الشكل (5) أن متوسطات قيم المادة العضوية للارتفاعات بين (400 - 500م) و (< 500م) كانت أعلى من بقية الارتفاعات ولكافة سنوات الدراسة من 2003 - 2004 وحتى 2022 على التوالي، وربما يعود تأثير كميات المادة العضوية في ترب منطقة الدراسة بالارتفاع عن سطح البحر إلى عدد من الأسباب من أهمها وفرة الغطاء النباتي في المناطق المرتفعة والتي تتميز بهطول مطرية أعلى نسبياً مقارنة بالمناطق المنخفضة، إضافة لانخفاض درجة حرارة ترب المناطق المرتفعة نسبياً مقارنة ببقية المناطق الأمر الذي يخفض من شدة عمليات المعدنة في تلك الترب والمحافظة على نسب من المادة العضوية في تلك الترب أعلى نسبياً من بقية الترب الأخرى، كما ويلاحظ من الشكل (5) أن متوسطات قيم المادة العضوية ظلت مستقرة نوعاً ما في صفي الارتفاعات عن سطح البحر (400 - 500م) و (< 500م) من 2003 - 2004 وحتى 2016 لتعود بعدها متوسطات قيم المادة العضوي في صفي الارتفاع السابقين للزيادة من 2016 وحتى 2022 وربما يشير ذلك إلى كمية الهطولات المطرية الجيدة التي تخللت تلك السنوات والتي أدت إلى تحسن ملحوظ في التغطية النباتية الموسمية والتي عملت على زيادة نسب المادة العضوية في طبقة التربة السطحية، في حين لم تحدث زيادة واضحة في متوسطات قيم المادة العضوية ضمن منطقة الدراسة

ذات الارتفاعات (200 - 300م) و (300 - 400م) وربما يعود ذلك لأسباب مختلفة منها الرعي الجائر وعودة النشاط البشري إلى تلك المناطق مقارنة بغيرها من مناطق الدراسة، إضافة إلى بعض المتغيرات المناخية التي تساعد على التمدن السريع للمخلفات النباتية وتناقص كميات المادة العضوية في التربة مقارنة بغيرها من مناطق الدراسة الأخرى، كما وأن تلك المناطق كانت خلال الفترات الماضية عرضة لزحف الرمال منها إلى مناطق أخرى داخل أو خارج منطقة الدراسة، أو إلى انتقال الرمال إليها من خارج منطقة الدراسة من خلال تشكل العواصف الغبارية وهذا ما يساعد على انخفاض نسب المادة العضوية في ترب تلك المناطق مقارنة ببقية أجزاء منطقة الدراسة الأخرى. الشكل (6) التالي يبين التوزيع المكاني لقيم المادة العضوية خلال الفترات الزمنية من 2003 - 2004.



الشكل (6) يبين التوزيع المكاني لقيم المادة العضوية للعام 2004 - 2003

حيث يتبين من خلال الشكل (6) أن القيم المرتفعة نسبياً من المادة العضوية في ترب منطقة الدراسة تركزت في الأجزاء الغربية والجنوبية الغربية من منطقة الدراسة، في حين

كانت القيم المنخفضة في الأجزاء الوسطى والشمالية الشرقية من منطقة الدراسة وهذا ينسجم ووجود غطاء نباتي جيد في الاجزاء المرتفعة من منطقة الدراسة خلال العام مقارنة بالأجزاء المنخفضة نسبياً والمعرضة للنشاط البشري من رعي جائر وتحطيب، إضافة لتدهور الخصائص الفيزيائية للتربة السطحية.

الاستنتاجات:

تم في هذه الدراسة استخدام خوارزمية RFR بكفاءة عالية لتقدير قيم المادة العضوية في ترب منطقة الدراسة حيث بلغت قيمة معامل الارتباط بيرسون Pearson $R = 0.661$ Correlation بين قيم المادة العضوية المتنبئ بها Predicted OM وقيم المادة العضوية المقاسة حقلياً Actual OM في بيانات الاختبار Testing Points البالغ عددها 112 قيمة والتي لم تستخدم في عملية تدريب النموذج في خوارزمية RFR. كما بلغت قيمة جذر متوسط مربع الفروق بين القيم الحقيقية والمتوقعة للمادة العضوية في التربة $RMSE = 0.516$ ، كما يتبين من خلال الدراسة أن القيم المرتفعة نسبياً من المادة العضوية في ترب منطقة الدراسة تركزت في الأجزاء الغربية والجنوبية الغربية من منطقة الدراسة، في حين كانت القيم المنخفضة في الأجزاء الوسطى والشمالية الشرقية من منطقة الدراسة خلال العام 2003 - 2004 وهذا ينسجم ووجود غطاء نباتي جيد في الاجزاء المرتفعة من منطقة الدراسة خلال العام مقارنة بالأجزاء المنخفضة نسبياً والمعرضة للنشاط البشري من رعي جائر وتحطيب.

المراجع العلمية:

المراجع الأجنبية:

1. Alvarez, R., Steinbach, H.S., Bono, A. (2011). An Artificial Neural Network Approach for Predicting Soil Carbon Budget in Agroecosystems. Soil Sci. Soc. Am.
2. Ballabio, C. (2009). Spatial prediction of soil properties in temperate mountain regions using support vector regression. Geoderma.
3. Bie, S.W., Beckett, P.H.T. (1971). Quality control in soil survey: II. The costs of soil survey. J. Soil Sci.
4. Bilgili, A.V., van Es, H.M., Akbas, F., Durak, A., Hively, W.D. (2010). Visible-near infrared reflectance spectroscopy for assessment of soil properties in a semi-arid area of Turkey. J. Arid. Environ.
5. Bobrovsky, M., Komarov, A., Mikhailov, A., Khanina, L. (2010). Modelling dynamics of soil organic matter under different historical land-use management techniques in European Russia. Ecol. Model.
6. Bogunovic, I., Trevisani, S., Pereira, P., Vukadinovic, V. (2018). Mapping soil organic matter in the Baranja region (Croatia): Geological and anthropic forcing parameters. Sci. Total Environ.
7. Chen, D., Chang, N., Xiao, J., Zhou, Q., Wu, W. (2019). Mapping dynamics of soil organic matter in croplands with

- MODIS data and machine learning algorithms. *Sci. Total Environ.*
8. Chen, F., Kissel, D.E., West, L.T., Adkins, W. (2000). Field-Scale Mapping of Surface Soil Organic Carbon Using Remotely Sensed Imagery. *Soil Sci. Soc. Am. J.*
 9. Chen, G., Ge, Z. (2019). SVM-tree and SVM-forest algorithms for imbalanced fault classification in industrial processes. *IFAC J. Syst. Control.*
 10. Chen, G., Li, S., Knibbs, L.D., Hamm, N.A.S., Cao, W., Li, T., Guo, J., Ren, H., Abramson, M.J., Guo, Y. (2018). A machine learning method to estimate PM_{2.5} concentrations across China with remote sensing, meteorological and land use information. *Sci. Total Environ.*
 11. Cheng, X.-F.; Shi, X.; Yu, D.; Pan, X.; Wang, H.; Sun, W. Using GIS spatial distribution to predict soil organic carbon in subtropical China. *Pedosphere* 2004, 14, 425–431.
 12. Dai, F., Zhou, Q., Lv, Z., Wang, X., Liu, G. (2014). Spatial prediction of soil organic matter content integrating artificial neural network and ordinary kriging in Tibetan Plateau. *Ecol. Indic.*
 13. Deiss, L., Margenot, A.J., Culman, S.W., Demyan, M.S. (2020). Tuning support vector machines regression models improves prediction accuracy of soil properties in MIR spectroscopy. *Geoderma.*

14. Farooq, I., Bangroo, S.A., Bashir, O., Shah, T.I., Malik, A.A., Iqbal, A.M., Mahdi, S.S., Wani, O.A., Nazir, N., Biswas, A. (2022). Comparison of Random Forest and Kriging Models for Soil Organic Carbon Mapping in the Himalayan Region of Kashmir. *Land*.
15. Feng, N., Hongfen, Z., Rutian, B. (2016). Hyperspectral prediction of soil organic matter content in the Reclamation cropland of Coal Mining Areas in the Loess Plateau. *Sci. Agric. Sin.*
16. Fernandes, M.M.H., Coelho, A.P., Fernandes, C., Da Silva, M.F., Marta, C.C.D. (2019). Estimation of soil organic matter content by modeling with artificial neural networks. *Geoderma*.
17. Grimm, R., Behrens, T., Märker, M., Elsenbeer, H. (2008). Soil organic carbon concentrations and stocks on Barro Colorado Island—Digital soil mapping using Random Forests analysis. *Geoderma*.
18. Guo, P. T., Wu, W., Sheng, Q. K., Li, M. F., Liu, H. B., Wang, Z. Y. (2013). Prediction of soil organic matter using artificial neural network and topographic indicators in hilly areas. *Nutr. Cycl. Agroecosyst.*
19. Heung, B., Bulmer, C.E., Schmidt, M.G. (2014). Predictive soil parent material mapping at a regional-scale: A Random Forest approach. *Geoderma*.

20. Kumar, S., Lal, R. (2011). Mapping the organic carbon stocks of surface soils using local spatial interpolator. *J. Environ. Monit.*
21. Kumar, S., Lal, R. (2011). Mapping the organic carbon stocks of surface soils using local spatial interpolator. *J. Environ. Monit.*
22. Kuter, S. (2021). Completing the machine learning saga in fractional snow cover estimation from MODIS Terra reflectance data: Random forests versus support vector regression. *Remote Sens. Environ.*
23. Lagacherie, P. (2008). Digital Soil Mapping: A State of the Art. *Digit. Soil Mapp. Ltd. Data.*
24. Liu, H.J., Zhang, B., Liu, D.W., Wang, Z.M., Song, K.S., Yang, F. (2008). Study on Quantitatively Remote Sensing Typical Soils in Songnen Plain, Northeast China. *J. Remote Sens.*
25. Manlay, R.J., Feller, C., Swift, M. (2007). Historical evolution of soil organic matter concepts and their relationships with the fertility and sustainability of cropping systems. *Agric. Ecosyst. Environ.*
26. Matheron, G. (1963). Principles of geostatistics. *Econ. Geol.*
27. Mccarty, G.W., Reeves, J.B., Reeves, V.B., Follett, R.F., Kimble, J.M. (2002). Mid-Infrared and Near-Infrared

- Diffuse Reflectance Spectroscopy for Soil Carbon Measurement. *Soil Sci. Soc. Am. J.*
28. Meersmans, J., De Ridder, F., Canters, F., De Baets, S., Van Molle, M. (2008). A multiple regression approach to assess the spatial distribution of Soil Organic Carbon (SOC) at the regional scale (Flanders, Belgium). *Geoderma*.
29. Meng, X., Bao, Y., Liu, J., Liu, H., Zhang, X., Zhang, Y., Wang, P., Tang, H., Kong, F. (2020). Regional soil organic carbon prediction model based on a discrete wavelet analysis of hyperspectral satellite data. *Int. J. Appl. Earth Obs. Geoinf.*
30. Menze, B.H., Kelm, B.M., Masuch, R., Himmelreich, U., Bachert, P., Petrich, W., Hamprecht, F.A. (2009). A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC Bio inform.*
31. Mishra, U., Torn, M.S., Masanet, E., Ogle, S.M. (2012). Improving regional soil carbon inventories: Combining the IPCC carbon inventory method with regression kriging. *Geoderma*.
32. Nocita, M., Stevens, A., Noon, C., van Wesemael, B. (2013). Prediction of soil organic carbon for different levels of soil moisture using Vis-NIR spectroscopy. *Geoderma*.

33. Pham, T.G., Kappas, M., Van Huynh, C., Nguyen, L.H.K. (2019). Application of ordinary kriging and regression kriging method for soil properties mapping in hilly region of Central Vietnam. *ISPRS Int. J. Geo-Inf.*
34. Qi, Y., Wang, Y., Chen, Y., Liu, J., Zhang, L. (2017). Soil organic matter prediction based on remote sensing data and random forest model in Shaanxi Province. *J. Nat. Resour.*
35. R Core Team. R. (2013). *A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria.*
36. Rosero-Vlasova, O.A., Vlassova, L., Pérez-Cabello, F., Montorio, R., Nadal-Romero, E. (2018). Modeling soil organic matter and texture from satellite data in areas affected by wildfires and cropland abandonment in Aragón, Northern Spain. *J. Appl. Remote Sens.*
37. Sparling, G.P., Wheeler, D., Vesely, E. T., Schipper, L.A. (2006). What is Soil Organic Matter Worth? *J. Environ. Qual.*
38. Takata, Y., Funakawa, S., Akshalov, K., Ishida, N., Kosaki, T. (2007). Spatial prediction of soil organic matter in northern Kazakhstan based on topographic and vegetation information. *Soil Sci. Plant Nutr.*
39. Van Raij, B., Andrade, J.C., de Cantarella, H., Quaggio, J.A. (2001). *Análise Química para Avaliação da Fertilidade*

- de Solos Tropicais; Instituto Agronômico: Campinas, Brazil.
40. Walkley, A., Black, I.A. (1934). An examination of the Degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method. *Soil Sci.*
41. Ward, K.J., Chabrilat, S., Neumann, C., Foerster, S. (2019). A remote sensing adapted approach for soil organic carbon prediction based on the spectrally clustered LUCAS soil database. *Geoderma.*
42. Were, K., Bui, D.T., Dick, Ø.B., Singh, B.R. (2015). A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape. *Ecol. Indic.*
43. Were, K., Bui, D.T., Dick, Ø.B., Singh, B.R. (2015). A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape. *Ecol. Indic.*
44. Wiesmeier, M., Barthold, F., Blank, B., Kögel-Knabner, I. (2011). Digital mapping of soil organic matter stocks using Random Forest modeling in a semi-arid steppe ecosystem. *Plant Soil.*

45. Ye, Y.; Wu, Q., Huang, J.Z., Ng, M.K., Li, X. (2013). Stratified sampling for feature subspace selection in random forests for high dimensional data. *Pattern Recognit.*
46. Zhang, C., Ma, Y. (2012). *Ensemble Machine Learning: Methods and Applications*; Springer: Berlin, Germany.
47. Zhang, C., Ma, Y. (2012). *Ensemble Machine Learning: Methods and Applications*; Springer: Berlin, Germany.
48. Zhang, L., Liu, Y., Li, X., Huang, L., Yu, D., Shi, X., Chen, H., Xing, S. (2018). Effects of soil map scales on simulating soil organic carbon changes of upland soils in Eastern China. *Geoderma*.
49. Zhang, M., Zhang, M., Yang, H., Jin, Y., Zhang, X., Liu, H. (2021). Mapping Regional Soil Organic Matter Based on Sentinel-2A and MODIS Imagery Using Machine Learning Algorithms and Google Earth Engine. *Remote Sens.* <https://doi.org/10.3390/rs13152934>.
50. Zhao, Z., Ashraf, M.I., Meng, F. R. (2013). Model prediction of soil drainage classes over a large area using a limited number of field samples: A case study in the province of Nova Scotia, Canada. *Can. J. Soil Sci.*
51. Zhao, Z., Yang, Q., Benoy, G., Chow, T.L., Xing, Z., Rees, H.W., Meng, F. R. (2010). Using artificial neural network models to produce soil organic carbon content distribution maps across landscapes. *Can.*

52. Zhao, Z., Yang, Q., Sun, D., Ding, X., Meng, F. R. (2020).
Extended model prediction of high-resolution soil organic
matter over a large area using limited number of field
samples. *Comput. Electron. Agric.*

Estimation and Studying of Soil Organic Matter in a Selected Area of the Syrian Desert Using Remote Sensing and Field Data

By

Omar Abdulrazzak ⁽¹⁾, Hani Ibrahim ⁽²⁾ , Wafa' Al-Hefel⁽³⁾

Professor in the Department of Soil Science - College of Agricultural ⁽¹⁾
Engineering - Al-Furat University

Researcher in General Organization of Remote Sensing – Damascus ⁽²⁾
– Syria.

PhD student in the Department of Soil Science - College of ⁽³⁾
Agricultural Engineering - Al-Furat University

ABSTRACT

The accurate understanding of spatial distribution and changes occurring in the surface soil organic matter is essential for land use planning and other activities related to agricultural management and environmental protection. In this study, LANDSAT 8 and LANDSAT 5 images were used to derive a set of spectral indices. Additionally, a digital elevation model was employed to derive relevant topographic indices related to soil organic matter. A large dataset of field measurements for organic matter values, synchronized with satellite image dates, and using Random Forest Algorithm RFA for determining the spatial distribution of organic matter in the soil of the study area. This area is part of the Syrian desert and spans between the Deir ez-Zor and Homs governorates. The study investigated changes in surface soil organic matter in the study area from 2003 to 2022, with a three-year time interval. The study demonstrated the effectiveness of using RFA and various remote sensing data to predict the amount of soil organic matter ratios over different time periods, thus assessing changes in surface soil organic matter within the studied region..

Keywords: Organic Matter, Satellite images, Random Forest, Syrian desert, Remote sensing.