



جمهوری اسلامی ایران



وزارت جهاد کشاورزی

سازمان تحقیقات، آموزش و ترویج کشاورزی

موسسه تحقیقات خاک و آب کشور



## نقشه شوری خاک‌های کشاورزی ایران

تهیه کننده

سعید سعادت

عضو هیات علمی موسسه تحقیقات خاک و آب کشور

همکاران

حامد رضایی<sup>1</sup>، لیلا اسماعیل نژاد<sup>1</sup>، رسول میرخانی<sup>1</sup> و یوسف رضا باقری<sup>2</sup>

<sup>1</sup>اعضای هیات علمی موسسه تحقیقات خاک و آب کشور

<sup>2</sup>محقق موسسه تحقیقات خاک و آب کشور

نشریه فنی: 630

1402

---

### مشخصات اثر

---

عنوان: نقشه شوری خاک‌های کشاورزی ایران

تهیه‌کنندگان: سعید سعادت، حامد رضایی، لیلا اسماعیل نژاد، رسول میرخانی و یوسف رضا باقری

ناشر: موسسه تحقیقات خاک و آب کشور

کارشناس انتشارات: سمانه پورمنصور

سال انتشار: 1402

حق چاپ برای ناشر محفوظ است.

این اثر با شماره 64447 در تاریخ 1402/8/27 در مرکز اطلاعات و مدارک علمی کشاورزی به ثبت رسیده است.

نقل مطالب با ذکر منبع بلامانع است.

---

نشانی: کرج، میدان استانداری، جاده مشکین‌دشت، بلوار امام خمینی (ره)، موسسه تحقیقات خاک و آب کشور

صندوق پستی: 311-31785

کد پستی: 3177993545

تلفن: 026-36201900

نمابر: 02636210121

پست الکترونیکی: info@swri.ir

وبسایت: http://www.swri.ir

---

مسئولیت صحت مطالب به عهده تهیه‌کنندگان است.

## فهرست مطالب

صفحه	عنوان
1.....	1- چکیده.....
2.....	2- اهمیت و ضرورت تهیه نقشه شوری خاک‌های کشاورزی ایران.....
3.....	3- سوابق تحقیق در داخل و خارج از کشور با تأکید بر نتایج آنها.....
6.....	4- مواد و روشها.....
8.....	4-2- انتخاب متغیرها.....
8.....	4-3- آنالیز رگرسیون و مدلسازی.....
8.....	4-3-1- روش bootstrap.....
11.....	4-3-2- درخت‌های طبقه‌بندی و رگرسیون.....
13.....	4-3-3- بسته‌بندی (Bagging).....
14.....	4-3-4- جنگل تصادفی (Random Forest).....
15.....	4-3-5- درخت رگرسیون تقویت‌کننده گرادیان.....
16.....	5- نتایج.....
18.....	ملاحظات.....
20.....	6- منابع مورد استفاده.....



## 1- چکیده

بر پایه گزارش فائو (2018)، سطح کل خاک‌های متأثر از نمک در جهان حدود 1130 میلیون هکتار است که 60 درصد آن خاک‌های شور، 26 درصد خاک‌های سدیمی، و 14 درصد نیز خاک‌های شور-سدیمی هستند. منطقه خاور میانه نیز با دارا بودن 176 میلیون هکتار، بیشترین سطح خاک‌های متأثر از نمک در جهان را دارا می‌باشد. وسعت اراضی تحت کشت آبی جهان هر سال حدود یک تا دو درصد کاهش می‌یابد (ICARDA, 2018). از طرفی، جمعیت جهان به سرعت در حال افزایش است و تا سال 2050 به حدود 10 میلیارد نفر خواهد رسید، بر این اساس، تولیدات غذای جهانی باید تا سال‌های 2025 و 2050 به ترتیب 38 و 57 درصد افزایش یابد. این در حالی است که در حال حاضر یک پنجم اراضی آبی تحت تأثیر نمک قرار دارند و سالانه 1/5 میلیون هکتار از اراضی به دلیل شوری بالا برای تولید محصولات کشاورزی نامناسب می‌شوند. اگر شور شدن خاک‌ها به همین منوال ادامه یابد تا سال 2050، 50 درصد از زمین‌های قابل کشت جهان از بین می‌رود. بنابراین لازم است تدابیر مدیریتی لازم برای حفظ امنیت غذایی اتخاذ شود. مدیریت منابع خاک شور نیازمند داشتن اطلاعات دقیق از وضعیت شوری اراضی کشاورزی، درجه شوری خاک‌ها، پراکندگی، وسعت و کلاس شوری در هر منطقه و در کاربری‌های مختلف کشاورزی (زرعی، باغی و آیش) می‌باشد. نیل به اهداف فوق تنها با وجود یک نقشه شوری دقیق و به روز امکان‌پذیر است. بر این اساس و با توجه به تغییرپذیری شوری خاک، با استفاده از داده‌های حدود 3220 پایگاه پایش کیفیت خاک‌های کشاورزی (سعادت و همکاران، 1398) نقشه شوری خاک‌های کشاورزی ایران تهیه شد. برای تهیه نقشه رقومی شوری خاک، افزون بر مقادیر شوری اندازه‌گیری شده، از داده‌های کمکی شامل داده‌های خاک، اقلیم، مدل رقومی ارتفاع، شاخص‌ها و باندهای مستخرج از تصاویر ماهواره‌ای سنتینل 2، نیز استفاده شد. همه مراحل پیش‌پردازش و پردازش تصاویر ماهواره‌ای در محیط گوگل ارث انجین و مدل‌سازی شوری خاک با استفاده از بسته نرم‌افزاری R 4.0.2 انجام شد. سایر مراحل آماده‌سازی و تهیه نقشه شوری نیز با استفاده از نرم‌افزار ARC GIS انجام گرفت. بر اساس نتایج

حاصل از تهیه این نقشه، حدود 58 درصد از اراضی کشاورزی کشور دارای شوری کمتر از دو دسی‌زیمنس بر متر بوده و در کلاس غیر شور قرار دارند. سطح اراضی کشاورزی با شوری کم (4-2 دسی‌زیمنس بر متر) حدود 14 درصد، با شوری متوسط (8-4 دسی‌زیمنس بر متر) 13/4 درصد، شوری زیاد (16-8 دسی‌زیمنس بر متر) 5/44 درصد، شوری نسبتاً زیاد (32-16 دسی‌زیمنس بر متر) 5/5 درصد و شوری خیلی زیاد (>32 دسی‌زیمنس بر متر) حدود 3/7 درصد می‌باشد.

## 2- اهمیت و ضرورت تهیه نقشه شوری خاک‌های کشاورزی ایران

در حال حاضر، شور شدن خاک یکی از مهم‌ترین محدودیت‌های استفاده از منابع خاک و آب بوده و عمدتاً توسط عوامل ثانویه و فعالیت‌های بشری، در حال گسترش می‌باشد. بررسی تغییرات شوری در خاک امری بسیار پیچیده بوده و تغییرات زمانی و مکانی آن به دلیل حجم آب مصرفی متفاوت، حرکت آب (نفوذپذیری و زهکشی) و تبخیر و تعرق متفاوت می‌باشد. بر اساس گزارش فائو، کل اراضی کشاورزی متأثر از نمک در جهان حدود 1130 میلیون هکتار است که 67 درصد آن متأثر از فعالیت‌های بشری است (FAO, 2018). این مقدار در ایران با دارا بودن اقلیم خشک و نیمه‌خشک و بر اساس اطلاعات موجود حدود 44/5 میلیون هکتار گزارش شده است (بنایی، 1380). بر اساس مطالعات و بررسی‌های انجام شده و مستند به آمارهای رسمی، سطح اراضی کشاورزی کشور حدود 18/5 میلیون هکتار می‌باشد. از این مقدار، حدود 8 میلیون هکتار به صورت آبی بوده که حدود 6/8 میلیون هکتار آن با محدودیت‌های مختلف و عمدتاً شوری روبرو است و تنها حدود 1/2 میلیون هکتار آن بدون محدودیت است (مؤمنی، 1389؛ سعادت و همکاران، 1398). همان‌گونه که از اطلاعات موجود پیداست، آمار ارائه شده در اینجا عمدتاً مربوط به مطالعات گذشته است، در صورتی که شوری از ویژگی‌های پویای خاک بوده و با تغییر مدیریت، اقلیم، و نیز با گذشت زمان می‌تواند تغییر کند. لذا با توجه به پویا بودن شوری، بررسی و ارائه آمارهای جدید و به‌روز از مقدار، وسعت و شدت آن ضروری است و تصمیم‌گیری در خصوص مدیریت آن می‌بایست مبتنی بر اطلاعات به‌روز و دقیق باشد. ایکاردا در سال 2014 با استفاده از

داده‌های موجود در بانک خاک فائو و نیز برآورد برخی اطلاعات از داده‌های موجود، نقشه شوری خاک جهان را ارایه نمود. آنچه که در این نقشه جلب توجه می‌نمود این بود که بیشتر اراضی ایران در رده‌بندی اراضی شور و اراضی شور-سدیمی طبقه‌بندی شده بود. اطلاعات ارایه شده، افزون بر اینکه با واقعیت خاک‌های کشاورزان تطابق نداشت، حاصل برآورد وضعیت شوری خاک‌های ایران با استفاده از اطلاعات کشورهای مجاور بود و در واقع برای تهیه این نقشه، از هیچ داده میدانی و اندازه‌گیری شده در کشور استفاده نشده بود. در سال 2021 مجدداً فائو نقشه شوری خاک‌های دنیا را ارایه نمود، لیکن اطلاعاتی از ایران در این نقشه دیده نمی‌شود و وضعیت شوری خاک‌های ایران در این نقشه به صورت سفید (بدون داده) نشان داده شده است. نقشه حاضر افزون بر اینکه می‌تواند خلأ اطلاعاتی موجود در نقشه جهانی ارائه شده توسط فائو را پر کند، با استفاده از آخرین تکنولوژی‌های تهیه نقشه و نیز بالاترین استانداردهای تولید نقشه با تعداد داده بسیار بیشتر از حد استاندارد معمول فائو تهیه شده و می‌تواند به عنوان الگویی برای تهیه نقشه شوری کشورهای دیگر قرار گیرد. لذا اطلاعات به‌دست آمده از این نقشه که حاصل انجام پژوهشی ملی در 3220 پایگاه مطالعاتی در اراضی با کاربری‌های متفاوت زراعی، باغی و آیش و در شرایط دیم و آبی است، دارای ارزش ملی و بین‌المللی است. در این نوشتار سعی خواهد شد تا با استفاده از داده‌های میدانی، تصویری شفاف از مقدار، شدت و وسعت شوری منابع خاک کشاورزی کشور ارائه شود تا بتواند به‌عنوان ابزاری کارآمد در تحلیل درست وضعیت منابع خاک کشور، به مدیران در انتخاب راه حلی منطقی و علمی برای کاهش و یا رفع محدودیت‌ها کمک نماید.

### 3- سوابق تحقیق در داخل و خارج از کشور با تأکید بر نتایج آنها

در ایران، شوری یک مسئله فراگیر و محدودکننده تولید پایدار کشاورزی است به طوری که قسمت‌های زیادی از مناطق خشک و نیمه خشک کشور، به‌ویژه فلات مرکزی، دشت‌های ساحلی جنوب و دشت خوزستان، مبتلا به سطوح مختلف شوری می‌باشند. بر خلاف بزرگی و گستردگی مسئله شوری، تاکنون اطلاعات موثقی متکی بر داده‌های کشور شمول و جامع ارائه نشده است. اطلاعات پراکنده و معدودی در

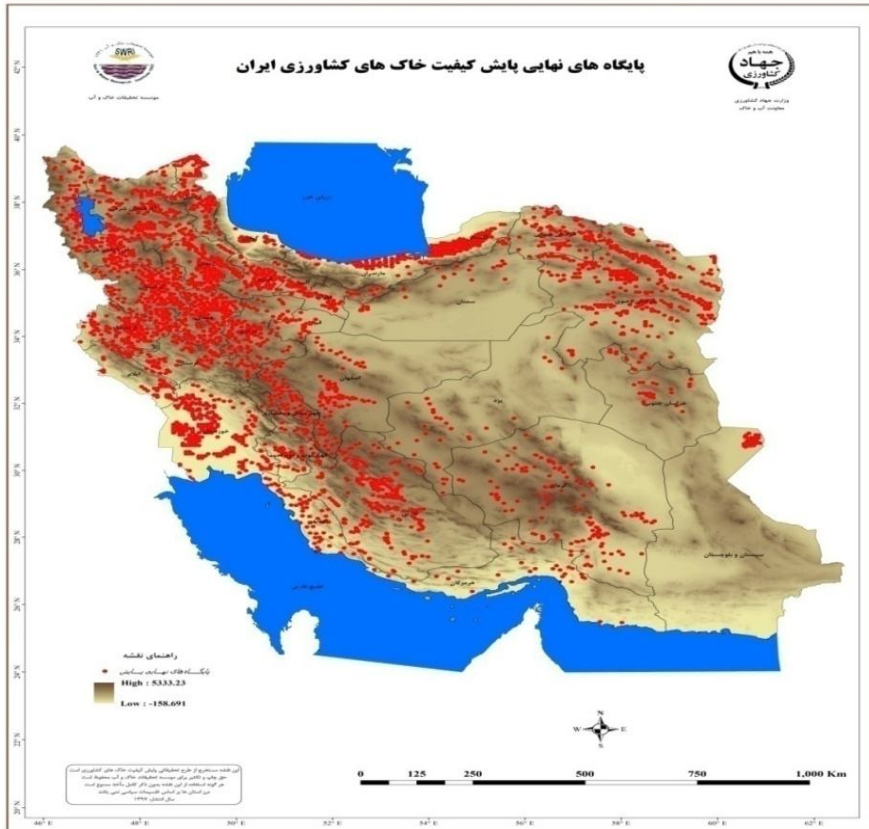
مورد وسعت و ویژگی‌های خاک‌های متأثر از نمک در ایران وجود دارند. در مطالعه انجام شده به‌وسیله (Dewan and Famouri, 1960) سطح اراضی متأثر از نمک در ایران حدود 15/5 میلیون هکتار برآورد شد که حدود 10 درصد از کل اراضی ایران را شامل می‌شد. در سال 1980، موسسه تحقیقات خاک و آب، سطح اراضی متأثر از نمک را 18 میلیون هکتار اعلام کرد (Soil and Water Research Institute, 1980). سیادت و همکاران (1990)، مساحت اراضی شور کشور را 16-23 میلیون هکتار بیان کردند که این تخمین شامل اراضی تحت کشت و بایر شامل هفت میلیون هکتار از مارش‌های نمکی دشت کویر و دشت لوت نیز بود. بر پایه داده‌های استخراج شده از نقشه خاک‌های ایران، خاک‌های با درجه شوری کم و متوسط حدود 25/5 میلیون هکتار و خاک‌های با درجه شوری زیاد حدود 8/5 میلیون هکتار را شامل می‌شد (FAO, 2000). مؤمنی (1389)، شوری منابع خاک کشور را با دو مقیاس سرزمین و اراضی کشاورزی مورد بررسی قرار داد. اطلاعات شوری منابع خاک در مقیاس سرزمین از نقشه یک میلیونیم منابع و استعداد خاک‌های ایران (بنائی، 1380) استخراج شد. نقشه فوق، بر اساس اطلاعات خاکشناسی موجود در آرشیو موسسه تحقیقات خاک و آب و اطلاعات خاکشناسی حاصل از انجام عملیات میدانی با سطوح دقت اجمالی، نیمه تفصیلی و تفصیلی توسط مهندسی مشاور و موسسه تحقیقات خاک و آب انجام شده بود تهیه شد. در مقیاس اراضی کشاورزی نیز، اطلاعات مربوط به 22 میلیون هکتار از اراضی کشاورزی که تا آن زمان دارای مطالعه خاکشناسی بود و مربوط به سال‌های 1332-1388 می‌شدند مورد بررسی قرار گرفتند. در مقیاس سرزمین، اراضی دارای خاک‌هایی با درجات مختلف شوری مساحتی معادل 55/6 میلیون هکتار (34 درصد مساحت کل کشور) را در برمی‌گرفتند. در مقیاس اراضی کشاورزی نیز حدود 6/8 میلیون هکتار مبتلا به شوری بودند. از این مقدار، حدود 4/3 میلیون هکتار جزو آن دسته از اراضی اعلام شدند که به غیر از شوری محدودیت دیگری ندارند. داده‌های حاصل از مطالعه مومنی (1389)، گرچه در برخی موارد خلا اطلاعاتی در مورد خاک‌های شور کشور را برطرف می‌کرد ولی مواردی نیز وجود داشت که استفاده از آن برای کاربران مشکل بود. به‌عنوان مثال، مناطقی روی نقشه وجود داشتند که کلاس شوری آنها بیان نشده بود یا مساحت کلاس‌های مختلف شوری قابل تشخیص نبود و در بسیاری از موارد، کلاس‌های شوری با



هم همپوشانی داشتند. از طرفی اطلاعات ارائه شده در آن مربوط به مطالعات انجام شده در یک بازه 56 ساله با مقیاس‌های متفاوت، روش‌های متفاوت و حتی زمان‌های نمونه‌برداری متفاوت بود. لذا با توجه به پویا بودن شوری خاک، نیاز به یک نقشه شوری به روز که اطلاعات آن مربوط به مطالعات جدید و انجام شده در یک دوره مشخص باشد کاملاً ضروری به نظر می‌رسید. از طرفی، بررسی سیر تاریخی مطالعات ذکر شده، موید این مطلب است که سطح خاک‌های شور کشور با گذشت زمان در حال تغییر است که بیان‌کننده مجموعه‌ای از اثرات تغییرات اقلیمی، مدیریت خاک، مدیریت آبیاری و کشت و کار بر تغییرات شوری خاک می‌باشد. لذا لازم است تا یک لایه اطلاعاتی رقومی برای تعیین سطح کلاس‌های شوری خاک، سطح خاک‌های شور در اراضی با کاربری‌های مختلف کشاورزی (زراعی، باغی و آیش) همچنین مساحت خاک‌های شور در هر استان یا منطقه تهیه شود. در بعد بین‌المللی نیز ارائه آمار و اطلاعات بعضاً اشتباه و بدون پشتوانه اطلاعات بومی از خاک‌های ایران، در سطوح بین‌المللی، منجر به ارایه اطلاعات غلط و تصمیم‌سازی‌های اشتباه در مدیریت خاک‌های شور کشاورزی ایران می‌شود. به عنوان مثال در سال 2014، ICARDA با استفاده از اطلاعات موجود در بانک اطلاعات جهانی خاک فائو، و با استفاده از روش‌های میان‌یابی مرسوم آن زمان، اقدام به تهیه نقشه خاک‌های شور- سدیمی سراسر جهان کرد. این در حالی بود که هیچ اطلاعات و آماری از ایران وجود نداشت و صرفاً بر پایه اطلاعات و آمار موجود از کشورهای همسایه، نقشه خاک‌های شور ایران ارایه شد. در این نقشه، سطح وسیعی از خاک‌های ایران در محدوده خاک‌های شور، سدیمی و شور-سدیمی قرار می‌گرفتند که با واقعیت‌های موجود خیلی تطابق نداشت. در سال 2021، فائو بر پایه یک رویکرد مشارکتی با همه کشورها، از آنها خواست یا اطلاعات و آمار شوری خاک‌ها را در اختیار فائو قرار دهند یا خود کشورها نقشه شوری خاک را تهیه نموده و در سایت فائو بارگزاری کنند. متأسفانه در این نقشه جهانی شوری خاک، هنوز اطلاعاتی از ایران ارائه نشده است. لذا لازم بود تا نسبت به تهیه نقشه شوری خاک‌های کشاورزی کشور اقدام شود.

#### 4- مواد و روش‌ها

داده‌ها و اطلاعات ارائه شده در این نوشتار، برگرفته شده از پژوهشی ملی و مبتنی بر روش‌های علمی با شبکه‌بندی کلیه اراضی کشاورزی کشور (شامل اراضی زراعی، باغی و آیش) در شرایط دیم و آبی می‌باشد. در این پژوهش، حدود 3200 پایگاه مطالعاتی در کل اراضی کشاورزی کشور (به ازاء حداقل هر 6000 هکتار یک پایگاه مطالعاتی) انتخاب گردید. در انتخاب پایگاه‌های مطالعاتی ترکیب متناسبی از اراضی زراعی، باغی و آیش در شرایط آبی و دیم در نظر گرفته شد و افزون بر این نقاط، نقاط و مکان‌های ویژه نظیر اراضی آلوده، شور، مانداب و ... که دارای سطح قابل توجهی در منطقه بودند نیز نمونه‌برداری شد. برای انتخاب این پایگاه‌های نمونه‌برداری، از اطلاعات جانبی دیگر نظیر نوع کاربری اراضی، نقشه خاک‌های ایران، اقلیم، اطلاعات زمین‌شناسی و غیره نیز استفاده شد. برای تعیین نقاط نمونه‌برداری، ابتدا با استفاده از اطلاعات موجود نظیر نقشه رقومی استان‌ها، نقشه رقومی کاربری اراضی و نیز نقشه خاک‌های کشور و با توجه به سطح زیر کشت استان‌ها (شامل زراعی، باغی و آیش) به تفکیک دیم و آبی، در یک شبکه 6×6 کیلومتر، پایگاه‌های پیشنهادی در هر منطقه مشخص شد. سپس با انجام بازدیدهای کارشناسی از هر یک از پایگاه‌ها و برگزاری جلسات کارشناسی با کارشناسان خبره منطقه‌ای نسبت به نهایی نمودن پایگاه‌ها اقدام گردید. با توجه به نوع کاربری و سطح کشاورزی منطقه، در نهایت تعداد 3220 پایگاه شامل 1656 پایگاه در اراضی زراعی آبی، 854 پایگاه در اراضی زراعی دیم، 392 پایگاه در اراضی باغی و 318 پایگاه در اراضی آیش انتخاب گردید. شکل یک موقعیت پایگاه‌ها را در کل کشور نشان می‌دهد. همان‌گونه که ملاحظه می‌شود پایگاه‌ها از پراکنش مناسبی برخوردار بوده و اطلاعات به‌دست آمده از آنها می‌تواند به کل منطقه/کاربری تعمیم یابد.



شکل 1- موقعیت پایگاه‌های مطالعاتی خاک

#### 1-4- داده‌ها و اطلاعات مورد استفاده در تهیه نقشه شوری

داده‌های مورد استفاده در تهیه نقشه شوری شامل دو دسته داده‌های اندازه‌گیری شده خاک و داده‌های استخراج شده از تصاویر ماهواره‌ای سنتینل 2 می‌باشد. داده‌های اندازه‌گیری شده خاک شامل مقادیر هدایت الکتریکی خاک مربوط به 3220 پایگاه به‌عنوان داده اصلی و داده‌های کمکی شامل آهک، سیلت، کلر، گچ، شن، سدیم، بیکربنات، رس، کلسیم، کربن آلی، منیزیم، پهاش، و نسبت جذب سدیم بود. همچنین اطلاعات استخراج شده از مدل رقومی ارتفاع (DEM) با استفاده از نرم‌افزار SAGA GIS و

همچنین پارامترهای اقلیمی نیز به‌عنوان داده کمکی مورد استفاده قرار گرفتند. پیش‌پردازش و پردازش تصاویر ماهواره‌ای برای استخراج ویژگی‌های طیفی، محاسبه شاخص‌های شوری و پوشش گیاهی، و همچنین استفاده از ابزارهای تبدیل طیفی مثل مؤلفه اصلی و تبدیل تسلدکپ در محیط گوگل ارث انجین انجام شد.

## 2-4- انتخاب متغیرها

از آنجا که داده‌های استفاده شده در مدلسازی، بویژه داده‌های طیفی دارای همبستگی بالایی با یکدیگر می‌باشند، با استفاده از الگوریتم‌های مناسب مثل تحلیل مؤلفه‌های اصلی و الگوریتم ژنتیک اقدام به انتخاب متغیرهای اصلی و مؤثر در برآورد شوری خاک شد.

## 3-4- آنالیز رگرسیون و مدلسازی

### 1-3-4- روش bootstrap

یکی از مهم‌ترین تکنیک‌های کاربردی و ناپارامتری در تخمین خطای مدل‌ها و محاسبات آماری، به‌کارگیری نمونه‌گیری مجدد (Resampling) و محاسبه برآوردگر تحت نمونه‌های مختلف است. در این حالت انحراف معیار حاصل از نمونه‌های مختلف، به عنوان خطای برآوردکننده شناخته می‌شود. به بیان ساده‌تر، چنین مکانیزمی برای پیدا کردن خطاهای برآوردکننده را بوت استرپ (Bootstrap) می‌نامند.

روش بوت استرپ یک روش مونت کارلو غیرپارامتریک است. ماهیت آن نمونه‌برداری مجدد از اطلاعات مشاهده‌ای و سپس استنتاج آماری در مورد ویژگی‌های توزیع جمعیت زیربنایی است (دانلیسون و همکاران، 2001). استراتژی بوت استرپ برای مقابله با نمونه‌های کوچک یا نمونه‌های توزیع نابرابر استفاده می‌شود. افزون بر این، روش بوت استرپ برای جلوگیری از مشکل کاهش نمونه ناشی از اعتبارسنجی متقابل، نمونه‌گیری مجدد انجام می‌دهد (وانگ و همکاران، 2020). از این روش می‌توان برای ایجاد دسته نمونه‌های تصادفی در داده‌ها نیز استفاده کرد. این فرآیند شامل سه مرحله است: (1) جایگزینی

نمونه‌ها. (2) محاسبه توزیع بوت‌استرپ و (3) تکرار مراحل بالا (N بار). برای استفاده از روش بوت‌استرپ از الگوریتم‌های یادگیری گروهی در تابع "BOOT" در R استفاده شد.

بوت‌استرپ در علم داده‌ها راهکاری قدرتمند است که برای استنباط آماری بدون تاکید بر فرضیه‌های زیاد استفاده می‌شود. راهکار فوق اجازه می‌دهد از یک نمونه داده، یک توزیع نمونه‌برداری ایجاد شود. برای استنباط آماری در بوت‌استرپ ضرورتی ندارد تا اطلاعاتی از توزیع آماری برآوردگر در اختیار داشته باشیم. در نتیجه بوت‌استرپ برای ایجاد فاصله اطمینان (Confidence Interval)، مدل رگرسیون (Regression Model) و یادگیری ماشین به شکل گسترده‌ای استفاده می‌شود. در این قسمت، مثالی در ارتباط با دو موضوع مرتبط با بوت‌استرپ در علم داده‌ها یعنی ساخت فاصله اطمینان و انجام آزمون فرض آماری ارائه خواهد شد. به عنوان مثال، کدهای دستوری برای تعیین میانگین شوری در اراضی دیم ایران ارائه شده است:

```
Salinity_all = pd.read_csv('Sal_all.csv')
Salinity_drylands = salinity_full.sample(n=3220)
[Salinity_dry['Sal_dry']==True].mean()
```

---Output---

2.99

اکنون بر مبنای کد زیر از تکنیک بوت‌استرپ در علم داده‌ها استفاده شده و یک فاصله اطمینان برای میانگین ایجاد می‌شود.

```
# Let's first create an empty list for storing means of bootstrapped samples
boot_means = []
# Let's write a loop for creating 1000 bootstrapped samples
for i in range(1000):
    boot_sample = Sal_dry.sample(200, replace=True)
    boot_sample_mean =
    boot_sample.[boot_sample['Salinity_drylands']==True]['height'].mean()
    boot_means.append(boot_sample_mean)
boot_means = np.array(boot_means)
```

برای حصول اطمینان 95 درصد یا در حقیقت 5 درصد خطا بر اساس 100 بار نمونه‌گیری، میانگین تخمین زده می‌شود. از بین آن‌ها 2/5 درصد از مقادیری که کمترین و بیشترین مقدار را دارند حذف، و میانگین‌های باقی‌مانده یک فاصله اطمینانی با ضریب 95 درصد را ایجاد می‌کنند. به بیان ساده‌تر کران پایین و بالای این میانگین‌ها همان فاصله اطمینان 95 درصدی برای میانگین جامعه آماری است.

```
# we build 95% in the middle portion
```

```
np.percentile(boot_means, 2.5), np.percentile(boot_means, 97.5)
```

```
---Output---
```

```
(66.00, 67.59)
```

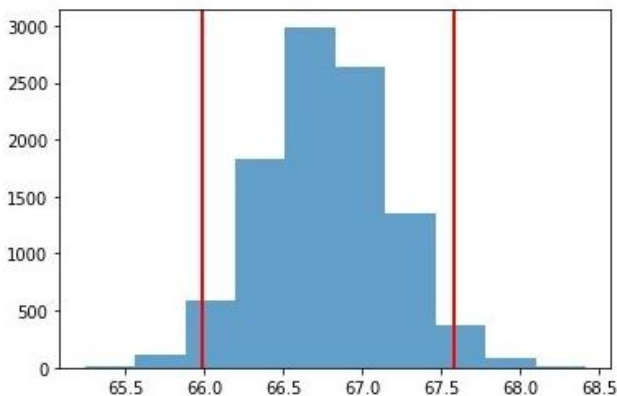
فاصله اطمینانی که در کد بالا ایجاد شده را می‌توان کران‌های میانگین برای شوری اراضی دیم با اطمینان 95 درصد توصیف کرد. برای نمایش این فاصله اطمینان از یک نمودار فراوانی استفاده می‌شود. کدی که در ادامه مشاهده می‌شود برای نمایش فراوانی میانگین‌های تولید شده توسط بوت استرپ است. ناحیه‌ای که مسافتی برابر با 95 درصد کل نمودار را به صورت متقارن از مرکز داده دارد همان فاصله اطمینان است.

```
plt.hist(boot_means, alpha=0.7);
```

```
plt.axvline(np.percentile(boot_means, 2.5), salinity=dry, linewidth=2) ;
```

```
plt.axvline(np.percentile(boot_means, 97.5), salinity=dry, linewidth=2) ;
```

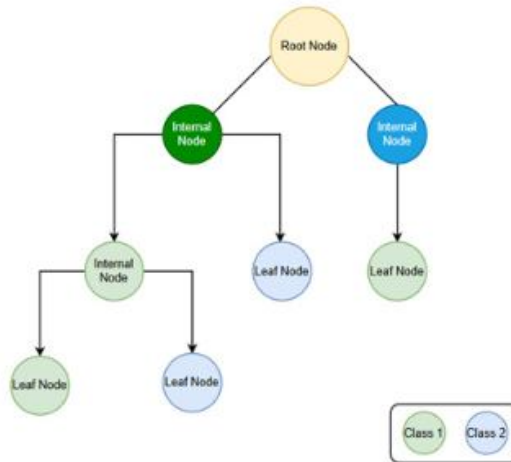
شکل زیر نمودار فراوانی حاصل را نشان می‌دهد.



شکل 2- نمودار فراوانی شوری در اراضی دیم حاصل از تکنیک بوت‌استرپ

## ۲-۳-۴- درخت‌های طبقه‌بندی و رگرسیون

الگوریتم CART نوعی الگوریتم طبقه‌بندی است که برای ساخت درخت تصمیم بر اساس شاخص ناخالصی جینی مورد نیاز است. این یک الگوریتم اساسی یادگیری ماشین است و طیف گسترده‌ای از موارد استفاده را ارائه می‌دهد. درخت تصمیم یک تکنیک است که برای تجزیه و تحلیل پیش‌بینی در زمینه‌های آمار، داده‌کاوی و یادگیری ماشین استفاده می‌شود. مدل پیش‌بینی در اینجا درخت تصمیم است و برای پیشرفت از مشاهدات در مورد نمونه‌هایی که توسط شاخه‌ها نشان داده می‌شود و در نهایت با مقدار هدف مورد نظر که در برگ‌ها نشان داده شده است، نتیجه‌گیری می‌شود. درختان تصمیم به دلیل خوانایی و سادگی، یکی از محبوب‌ترین روش‌های یادگیری ماشین هستند. ساختار یک درخت تصمیم از سه بخش اصلی تشکیل شده است: گره‌ها ریشه، گره‌های داخلی و گره‌های برگ.



شکل 3- شمایی از یک درخت تصمیم‌گیری

همان‌طور که در نمودار نشان داده شده است، اولین گره یا گره ریشه مجموعه داده‌های آموزشی است و بعد از آن گره داخلی و گره برگ. گره داخلی به‌عنوان گره تصمیم‌گیری عمل می‌کند، زیرا این نقطه‌ای است که در آن گره بر اساس بهترین ویژگی زیر گروه تقسیم می‌شود. گره نهایی یا گره برگ است که تصمیم را مشخص می‌کند.

الگوریتم CART در درخت تصمیم، گره‌ها را بر اساس مقدار آستانه یک ویژگی به زیرگره‌ها تقسیم می‌کند و این کار را با جستجوی بهترین ویژگی همگن برای زیرگره‌ها، با کمک معیار شاخص جینی انجام می‌دهد. گره ریشه به عنوان مجموعه آموزشی در نظر گرفته می‌شود و با در نظر گرفتن بهترین ویژگی و مقدار آستانه به دو قسمت تقسیم می‌شود. افزون بر این، زیر مجموعه‌ها نیز با استفاده از همان منطق تقسیم می‌شوند. این کار ادامه می‌یابد تا آخرین زیر مجموعه خالص در درخت یا حداکثر تعداد برگ ممکن در آن درخت در حال رشد یافت شود. این به هرس درخت نیز معروف است. شاخص جینی به شرح زیر تعریف می‌شود:

$$GI = \sum_{i=0}^c P_i(1 - P_i)$$

Which can be written as:

$$GI = 1 - \sum_{i=0}^c P_i^2$$

الگوریتم درختان طبقه‌بندی و رگرسیون (CART) زیر مجموعه جنگل تصادفی است که یکی از قوی‌ترین الگوریتم‌های یادگیری ماشین است. الگوریتم CART به‌عنوان مجموعه‌ای از داده‌هایی است که بر اساس ویژگی‌هایشان سازمان‌دهی می‌شوند. این الگوریتم به‌طور گسترده‌ای در تصمیم‌گیری درختان از طریق طبقه‌بندی و رگرسیون استفاده می‌شود. همچنین به‌طور گسترده‌ای در داده‌کاوی برای ایجاد مدلی که ارزش یک هدف را بر اساس مقادیر بسیاری از متغیرهای ورودی (یا متغیرهای مستقل) پیش‌بینی می‌کند، استفاده می‌شود. مدل CART تحت بسته "rpart" در داخل R پردازش می‌شود. از مزایای این روش می‌توان به موارد زیر اشاره نمود:

- الگوریتم CART غیر پارامتری است، بنابراین به اطلاعات مربوط به نوع خاصی از توزیع بستگی ندارد.
- الگوریتم CART هر دو آزمایش را با مجموعه داده‌های آزمایش و اعتبار سنجی متقابل ترکیب می‌کند تا میزان تناسب را با دقت بیشتری اندازه‌گیری کند.

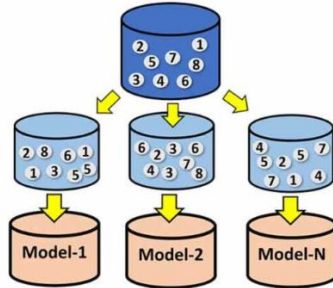


- به کاربر این امکان را می‌دهد که چندین بار از متغیرهای یکسان در مناطق مختلف درخت استفاده کند. این مهارت قادر است وابستگی متقابل پیچیدگی بین گروه‌های متغیر را آشکار کند.
- داده‌های پرت در متغیرهای ورودی تأثیر معنی‌داری بر CART ندارند.
- می‌توان محدودیت‌های توقف را لغو کرد تا درختان تصمیم‌گیری بیش از حد رشد کنند و سپس درخت را به اندازه ایده آل خود برش دهند. این روش با توقف زود هنگام، احتمال از دست رفتن ساختار اساسی در مجموعه داده‌ها را کاهش می‌دهد. برای انتخاب مجموعه ورودی متغیرها، می‌توان از CART در ترکیب با سایر الگوریتم‌های پیش‌بینی استفاده کرد.
- الگوریتم CART می‌تواند بسیار مقاوم باشند. یک تغییر کوچک در داده‌های آموزش می‌تواند منجر به تغییر بزرگی در درخت و در نتیجه پیش‌بینی‌های نهایی شود.

### ۳-۳-۴ - بسته‌بندی (Bagging)

تکنیک Bagging در سال 1996 توسط Breiman مطرح شد. Bagging یک تکنیک یادگیری جمعی است که هدف آن کاهش خطای یادگیری با استفاده از مجموعه‌ای از مدل‌های یادگیری ماشین هم‌نوع/همگن هست. منظور از هم‌نوع یعنی همه مدل‌های پایه در این تکنیک جنس یکسانی دارند. برای مثال همه مدل‌ها درخت تصمیم هستند. در این تکنیک هدف کاهش واریانس است که با اینکار خطای طبقه‌بندی یا رگرسیون کاهش می‌یابد. از آنجا که تکنیک Bagging با هدف کاهش واریانس مدل ارائه شده است، به همین دلیل، مدل پایه باید پتانسیل *overfitting* بالایی داشته باشد، یعنی مدل پایه واریانس بالایی داشته باشد. در غیر این صورت تکنیک Bagging کارساز نخواهد بود. چون زمانی این تکنیک خوب عمل خواهد کرد که تنوع بین خروجی‌ها زیاد باشد. اگر مدل‌ها واریانس بالایی نداشته باشند تنوعی هم بین خروجی مدل‌ها نخواهد بود. لذا بهتر است مدل پایه را یک مدل غیرخطی انتخاب کنیم. چرا که مدل‌های غیرخطی احتمالاً *overfitting* بالایی دارند. مثل شبکه‌های عصبی چندلایه، درخت تصمیم و.... این تکنیک، با هدف کاهش واریانس، مسئله *overfitting* را حل نموده و باعث افزایش عملکرد نهایی مدل‌سازی می‌شود.

در این تکنیک در ابتدا نوع و تعداد مدل‌های پایه مشخص می‌شود، سپس برای هر مدل پایه، یک داده آموزشی از انتخاب شده و آموزش داده می‌شود.



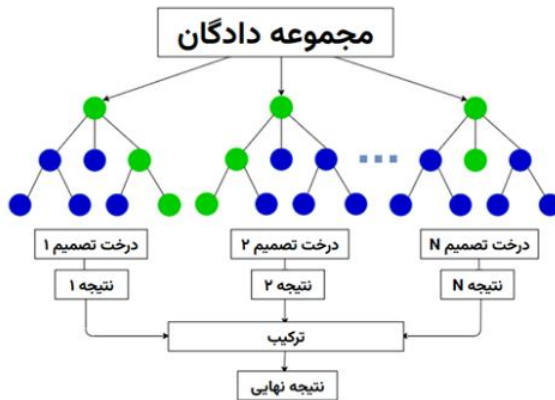
شکل 4- تصویر شماتیک از روش بسته بندی یا Bagging.

پروسه آموزش بسیار ساده است، در ابتدا چندین زیرمجموعه از روی داده آموزشی ساخته می‌شود، و سپس مدل‌های پایه با تک تک این مجموعه‌ها به صورت جداگانه آموزش داده می‌شود. جنس/ نوع مدل‌های پایه یکی است، اما از آنجا که هر کدام با یک مجموعه متفاوتی آموزش می‌بینند، در نتیجه دانش خروجی متفاوتی خواهد داشت و در پروسه تصمیم‌گیری می‌توانند نظر متفاوتی نسبت به هم برای یک ورودی خاص داشته باشند. الگوریتم BG با بسته "ipred" در R اجرا شد.

#### ۴-۳-۴ - جنگل تصادفی (Random Forest)

RF یک فناوری یادگیری گروهی است که از مدل CART مشتق شده است و به عنوان یادگیرنده ضعیف عمل می‌کند. هر RF از درختان زیادی تشکیل شده است که از نمونه های bootstrap برای مدلسازی استفاده می‌کنند (چن و همکاران، 2019b). برآورد خطای خارج از بسته (OOB, out of bag) امتیازی برای اندازه‌گیری خطای تعمیم یک مدل است. این اندازه‌گیری مشابه اعتبارسنجی متقاطع است. در هر گره، یک زیرمجموعه از همه ویژگی‌ها به طور تصادفی انتخاب شده تا بهترین روش تقسیم بندی محاسبه شود. در نهایت، نتایج همه یادگیرندگان ضعیف به عنوان خروجی مدل به طور میانگین محاسبه خواهد شد (هنگ و همکاران، 2019). در فرایند رگرسیون، RF دارای مزایایی است: (1)

با داده‌های با ابعاد بالا سروکار داشته و سطوح اهمیت ویژگی‌ها را نشان می‌دهد (2) نتایج چندین یادگیرنده ضعیف را برای بهبود دقت پیش‌بینی میانگین می‌گیرد. (3) محاسبات موازی می‌تواند به راحتی پیاده‌سازی شود. سه پارامتر زیر در مدل RF کلیدی هستند: تعداد درختان رگرسیون (ntree)، تعداد متغیرهایی که به‌طور تصادفی به‌عنوان کاندید در هر تقسیم نمونه‌گیری می‌شوند (mtry)، و حداقل اندازه یک گره پایانی (اندازه گره). مدل RF با بسته "RandomForest" در R انجام شد.



شکل 5- نمایش از روش جنگل تصادفی (Random Forest).

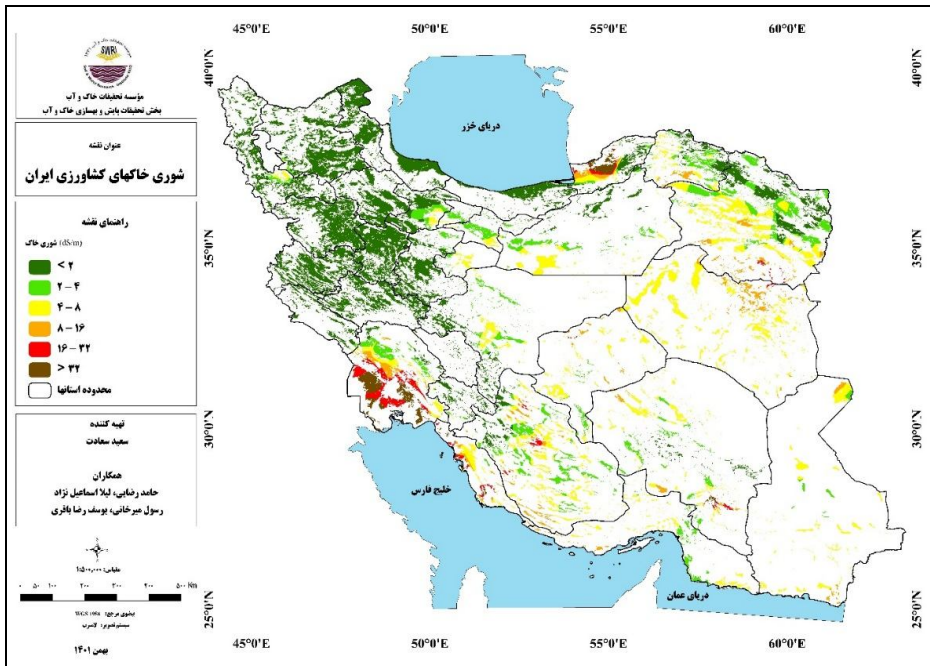
#### ۵-۳-۴- درخت رگرسیون تقویت‌کننده گرادیان

درخت رگرسیون تقویت‌کننده گرادیان (GBRT)، یک روش یادگیری آماری غیرپارامتریک بسیار مؤثر و انعطاف‌پذیر، با ترکیب درخت‌های رگرسیون متعدد بر اساس نظریه تقویت (Ridge-way, 2007) ساخته شده است. یک الگوریتم GBRT از تقویت گرادیان تصادفی استفاده می‌کند. (چن و همکاران، 2019a). بنابراین، یک GBRT می‌تواند استحکام بیشتری به مدل دهد. رده‌بندی تقویت گرادیان از روش ارتقاء استفاده می‌کند؛ این روش وام‌گرفته از روش درخت تصمیم است. یادگیرنده‌های ضعیف موجود بعد از دریافت وزن بیشتر، به درخت بعدی منتقل می‌شوند تا پیش‌بینی‌های هر درخت از

درخت قبلی بهتر باشد. خروجی، میانگین پیش‌بینی‌های نهایی است. الگوریتم تقویت گردیدن زمانی که داده‌های زیاد با پیش‌بینی‌های بالا داریم، مورد استفاده قرار می‌گیرند.

## 5- نتایج

نقشه شوری خاک‌های کشاورزی ایران در شکل شش آورده شده است. این نقشه نشان‌دهنده درجات شوری خاک در اراضی کشاورزی ایران با کاربری‌های مختلف زراعی، باغی و آیش به دو صورت دیم و آبی می‌باشد. بر این اساس، مناطق شمال غربی، همچنین اراضی کشاورزی واقع در دامنه‌های زاگرس، بیشتر در کلاس بدون شوری (هدایت الکتریکی کمتر از دو دسی‌زیمنس بر متر) و یا دارای شوری کم (هدایت الکتریکی 4-2 دسی‌زیمنس بر متر) هستند. خاک‌های با شوری‌های بیشتر در مناطق مرکزی و جنوبی کشور واقع شده‌اند.



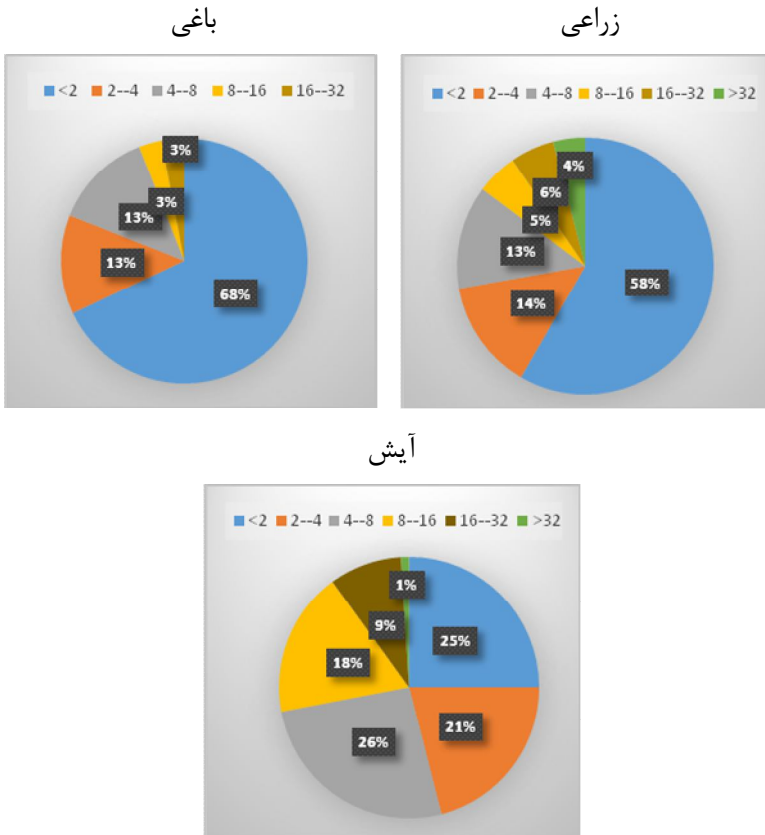
شکل 6- نقشه شوری خاک‌های کشاورزی ایران

بر پایه جدول یک، حدود 58 درصد از اراضی کشاورزی کشور دارای شوری کمتر از دو دسی‌زیمنس بر متر بوده و در کلاس غیر شور قرار دارند. سطح اراضی کشاورزی با شوری کم (2-4 دسی‌زیمنس بر متر) حدود 14 درصد، با شوری متوسط (8-4 دسی‌زیمنس بر متر) 13/4 درصد، شوری زیاد (8-16 دسی‌زیمنس بر متر) 5/44 درصد، شوری نسبتاً زیاد (16-32 دسی‌زیمنس بر متر) 5/5 درصد و شوری خیلی زیاد ( $>32$  دسی‌زیمنس بر متر) حدود 3/7 درصد می‌باشد.

جدول 1- سطح اراضی کشاورزی با درجات مختلف شوری خاک

سطح (درصد)	کلاس شوری
57/96	<2
14	2-4
13/4	4-8
5/44	8-16
5/5	16-32
3/7	>32
100	مجموع

شکل هفت، سطح هر کلاس شوری خاک در اراضی زراعی با کاربری‌های مختلف را نشان می‌دهد. بر این اساس، 58 درصد اراضی در کاربری‌های زراعی غیر شور بوده و بیشترین سطح شوری به‌ترتیب در کلاس 2-4 و 4-8 دسی‌زیمنس بر متر واقع شده‌اند. در کاربری‌های باغی، 68 درصد اراضی غیر شور و بیشترین سطح اراضی شور در همان کلاس‌های 2-4 و 4-8 دسی‌زیمنس بر متر قرار دارند.



شکل 7- وضعیت شوری خاک در کاربری‌های مختلف

### ملاحظات

به لحاظ اهمیت ملی و بین‌المللی، این نقشه در تاریخ 1401/11/17 توسط وزیر محترم وقت جهاد کشاورزی با حضور معاونین محترم وزیر، رؤسای مؤسسات و مراکز پژوهشی و کارشناسان و محققان ذی‌ربط در محل سازمان تحقیقات، آموزش و ترویج کشاورزی رونمایی شد.



## 6- منابع مورد استفاده

1. بنائی، م. ح. 1380. نقشه منابع و استعداد خاک‌های ایران. مؤسسه تحقیقات خاک و آب، تهران، ایران، 6 برگ.
2. سازمان نقشه‌برداری کشور. 1398. اطلس آب، انتشارات سازمان نقشه‌برداری کشور.
3. سعادت، س. 1398. گزارش نهایی طرح پایش کیفیت خاک‌های کشاورزی. مؤسسه تحقیقات خاک و آب. سازمان تحقیقات، آموزش و ترویج کشاورزی.
4. مؤمنی، ع. 1389. پراکنش جغرافیایی و سطوح شوری منابع خاک ایران. مجله پژوهش‌های خاک (علوم خاک و آب): الف، جلد 24، شماره 3. صفحات 203-215.
5. وزارت نیرو. 1400. چاه‌های آب، فرونشست زمین و برنامه هفتم توسعه، پایگاه اطلاع‌رسانی وزارت نیرو (پاون).
6. Aksoy, S., A. Yildirim, T. Gorji, N. Hamzehpour, A. Tanik, and E. Sertel. 2022. Assessing the performance of machine learning algorithms for soil salinity mapping in Google Earth Engine platform using Sentinel-2A and Landsat-8 OLI data. *Advances in Space Research*
7. Ayers, R. S., and Westcot, D. W. (1985). *Water Quality for Agriculture*. Rome: FAO. FAO Irrigation and Drainage Paper no. 29.
8. Bauder, T. A., Waskom, R. M., Sutherland, P. L., and Davis, J. G. (2011). *Irrigation Water Quality Criteria, Crop Series/irrigation. Fact Sheet No. 0.506*. Fort Collins: Colorado State University Extension Publication.
9. FAO. 2020. Mapping of salt-affected soil, technical manual. Rome, FAO. <https://doi.org/10.4060/ca9215en>.
10. FAO. 2018. Handbook for saline soil management. Editors: R. Vargas, E.I. Pankova, S.A. Balyuk, P.V. Krasilnikov and G.M. Khasankhanova. Published by the Food and Agriculture Organization of the United Nations and Lomonosov Moscow State University
11. Muhetaer, n., I. Nurmemet, A. Abulaiti, S. Xiao, and J. Zhao. 2022. A Quantifying Approach to Soil Salinity Based on a Radar Feature Space Model Using ALOS PALSAR-2 Data. *Remote Sensing*. Volume 14, issue 2.
12. Saadat, S., Dehghany, F., Rezaei, H., Esmaelnejad, L., and Maleki, P. 2020. Effects of salinity and Ca:Mg ratio of irrigation water on pistachio seedlings phosphorus planted under greenhouse conditions. *Desert* (25):1, 25-32
13. Singh, A. 2022. Soil salinity: a Global Threat to Sustainable Development. *Soil use Manage.* 38:39-67.



- 
14. Soil Survey Staff (1954). Diagnosis and Improvement of saline and Alkali Soils. Agriculture Handbook, U.S. Department of Agriculture. Nat. Resour. Conservation Serv. 60, 83–100. doi:10.1093/aibsbulletin/4.3.14-a
  15. Stavi I, Thevs N and Priori S. 2021. Soil Salinity and Sodicity in Drylands: A Review of Causes, Effects, Monitoring, and Restoration Measures. *Front. Environ. Sci.* 9:712831. doi: 10.3389/fenvs.2021.712831 .Volume 69, Issue 2, Pages 1072-1086
  16. Wilson, M.J. 2019. The importance of parent material in soil classification: A review in a historical context. *Catena*, 182, doi: 10.1016/j.catena.2019.104131.