

مقایسه روش‌های نمونه‌برداری k -NN در برآورد تراکم درختان بنه (*Pistacia atlantica* Desf.) با الگوی مکانی کپه‌ای در یک توده تنک زاگرس

معصومه موصول^۱ و سید یوسف عرفانی فرد^{۲*}

۱- دانشجوی کارشناسی ارشد، دانشکده کشاورزی، دانشگاه شیراز، شیراز، ایران

۲- نویسنده مسئول، استادیار، دانشکده کشاورزی، دانشگاه شیراز، شیراز، ایران. پست الکترونیک: erfani@shirazu.ac.ir

تاریخ پذیرش: ۹۴/۰۱/۳۱

تاریخ دریافت: ۹۳/۰۸/۰۳

چکیده

تراکم (تعداد درختان در واحد سطح) یکی از مشخصه‌های ساختاری مهم در توده‌های جنگلی است که در درک پویایی جنگل مناسب است. روش k امین نزدیکترین همسایه (k -NN) یک روش فاصله‌ای است که به‌طور متداول در آماربرداری جنگل برای برآورد مشخصه‌های کمی به‌کار می‌رود. در این مطالعه روش k -NN با پنج راهکار نزدیکترین فرد (NI)، نزدیکترین همسایه (NN)، جفت‌های تصادفی (RP)، چارک نقطه مرکز (PCQ) و همسایه چارکی (QN) برای برآورد تراکم درختان بنه (*Pistacia atlantica* Desf.) در جنگل‌های زاگرس استفاده شد. یک توده با مساحت ۴۵ هکتار در رایشگاه تحقیقاتی بنه انتخاب شد و به‌منظور تعیین تراکم واقعی، آماربرداری صددرصد شد. توزیع مکانی درختان بنه، کپه‌ای ($R=0.79$ و $\alpha=-12/38$) و تراکم آنها $19/44$ درخت در هکتار بود. سپس روش k -NN با راهکارهای مختلف و k برابر دو تا ۱۰ در ۴۶ نقطه نمونه‌برداری در یک شبکه 100×100 متری مورد استفاده قرار گرفت. نتایج نشان داد که به‌جز راهکار PCQ، روش k -NN با سایر راهکارها مقدار تراکم را به‌طور معنی‌داری (در سطح اطمینان ۹۵ درصد) درست برآورد کرد. تعداد k و راهکار انتخاب آن بر صحت و دقت نتایج روش k -NN تأثیر گذاشت، به‌نحوی که $k=4$ در NI، $k=7$ در NN و $k=2$ در RP و QN تراکم درختان بنه را با کمترین خطا (RMSE و اریبی) برآورد کردند. به‌طور کلی نتایج نشان داد که روش k -NN بهینه با k و راهکار مناسب توانسته تراکم درختان بنه با توزیع مکانی کپه‌ای را در یک توده تنک در جنگل‌های زاگرس برآورد کند.

واژه‌های کلیدی: بنه، تراکم، زاگرس، خطا، k -NN

مقدمه

جنگل‌شناسی، مطالعه ساختار و مدیریت جنگل تأثیر بسزایی دارد (Guebas & Koedam, 2006). روش صحیح اندازه‌گیری تراکم در یک توده، شمارش تعداد کل درختان و اندازه آن در واحد سطح توده مذکور است که به‌دلیل دشواری در اندازه‌گیری تعداد همه درختان، چندان متداول نیست (Chirici et al., 2011)، بنابراین دسترسی به یک روش مناسب و نارپی در برآورد تراکم درختان در جنگل

تراکم یکی از مهمترین ویژگی‌های ساختاری در هر جامعه جنگلی است که سایر ویژگی‌های جامعه مانند گونه، جنسیت و سن بر آن تأثیر فراوانی دارند و به همین دلیل ویژگی مناسبی در ارزیابی پویایی جوامع جنگلی محسوب می‌شود (Kangas & Maltamo, 2006; Bonham, 2013). برآورد صحیحی از تراکم در انتخاب روش‌های مناسب

ضروری است.

روش‌های متعددی برای برآورد تراکم درختان توسعه پیدا کرده‌اند که می‌توان آنها را به دو گروه اصلی قطعه‌نمونه-ای و فاصله‌ای تقسیم کرد. از بین این دو گروه، روش‌های فاصله‌ای به دلیل کاربردی بودن، هزینه کمتر و انعطاف‌پذیری بیشتر در شرایط مختلف، بیشتر مورد استفاده قرار گرفته‌اند (Magnussen *et al.*, 2012a; Bonham, 2013). صحت برآورد تراکم در جنگل با استفاده از روش‌های فاصله‌ای به برآوردکننده‌های مناسبی وابسته است که براساس فاصله بین نقطه نمونه‌برداری تا درخت و یا فاصله درختان از یکدیگر عمل می‌کنند (Hanberry *et al.*, 2012). یکی از روش‌های فاصله‌ای مهم در برآورد تراکم درختان در جنگل، روش k امین نزدیکترین همسایه (k -Nearest Neighbour: k -NN) است که در آن براساس فاصله تا k امین درخت، می‌توان تراکم را برآورد کرد. در این روش، همه قطعه‌نمونه‌ها تعداد k درخت مساوی دارند و بنابراین قطعه‌نمونه‌های خالی یا متراکم در آن وجود ندارند که منجر به اریبی نتایج شوند (Kleinn & Vilcko, 2006; Kung *et al.*, 2012). هرچند کاربرد روش k -NN در برآورد تراکم در جنگل در مطالعات داخل (Jahan *et al.*, 2010; Karamshahi *et al.*, 2012) و خارج از کشور (Zhu & Zhang, 2009; Hijbeek *et al.*, 2013) مورد تأیید قرار گرفته است، بررسی مقایسه‌ای راهکارهای متنوع پیاده‌سازی این روش که تاکنون کمتر مورد توجه قرار گرفته، ضروری به نظر می‌رسد.

برای استفاده عملی از روش k -NN در جنگل، امکان انتخاب k امین درخت در هر نقطه نمونه‌برداری با راهکارهای مختلفی امکان‌پذیر است که این موضوع بر صحت و دقت نتایج به دست آمده از این روش تأثیر فراوانی دارد (Kleinn & Vilcko, 2006). به طور کلی، پنج راهکار روش k -NN، نزدیکترین فرد (Nearest Individual: NI) (Engeman *et al.*, 1994)، نزدیکترین همسایه (Nearest Neighbour: NN) (Temesgen *et al.*, 2012)، جفت‌های تصادفی (Random Pairs: RP) (Hill *et al.*, 2005)، چارک نقطه‌مرکز (Point-Centered

(Quarter: PCQ) (Bryant *et al.*, 2004; Hijbeek *et al.*, 2013) و همسایه چارکی (Quartered Neighbour: QN) (Zhu & Zhang, 2009; Leavitt & Clair, 2011) هستند که جداگانه در مطالعات مختلف مورد بررسی قرار گرفته‌اند. اگرچه در پژوهش‌های محدودی به مقایسه برخی از راهکارهای روش k -NN پرداخته شده است (به عنوان مثال راهکارهای PCQ و QN در Zhang و Zhu (۲۰۰۹) و راهکارهای NI و PCQ در Kleinn و Vilcko (۲۰۰۵))، مقایسه کارایی روش k -NN با راهکارهای مختلف، انتخاب راهکار بهینه این روش را در برآورد تراکم درختان ممکن می‌کند.

جنگل‌های زاگرس که به عنوان یکی از منابع بیولوژیک و ذخایر ژنتیک کشور هستند، از جنبه‌های بوم‌شناختی، اقتصادی و اجتماعی از اهمیت ویژه‌ای برخوردارند و پایش آنها با استفاده از ویژگی‌های ساختاری مانند تراکم، مدیریت پایدار این جنگل‌ها اهمیت بسزایی دارد. بنه (*Pistacia atlantica* Desf.) دومین گونه مهم در این جنگل‌ها پس از بلوط ایرانی است که علاوه بر ایفای نقش بوم‌شناختی در جنگل‌های زاگرس، به دلیل تولید محصولات فرعی ارزشمند (میوه و سقز) از اهمیت اقتصادی و اجتماعی زیادی برخوردار است (Jazirehi & Ebrahimi Rostaghi, 2003; Pourreza *et al.*, 2008).

آگاهی از راهکارهای متنوع انتخاب درختان در روش k -NN به دلیل تأثیر آنها بر کارایی روش مذکور در برآورد تراکم درختان در جنگل باید مورد توجه قرار گیرد تا امکان انتخاب راهکار بهینه روش مذکور و دستیابی به برآوردی ناریب از تراکم ممکن شود. بنابراین این پژوهش با هدف معرفی و کاربرد راهکارهای پیاده‌سازی روش k -NN در برآورد تراکم درختان بنه در بخشی از جنگل‌های زاگرس انجام شد. همچنین بررسی مقایسه‌ای هر یک از این راهکارها در شرایط مشابه (منطقه مورد مطالعه و روش نمونه‌برداری) بررسی تأثیر هر کدام را بر صحت برآورد تراکم توسط روش k -NN امکان‌پذیر می‌کند. در این پژوهش از برآوردکننده‌هایی استفاده شد که در مقایسه با انواع

آماربرداری صد درصد، تراکم واقعی درختان بنه مشخص شد که از آن در ارزیابی نتایج به دست آمده از روش k -NN استفاده شد. همچنین یک شبکه آماربرداری 100×100 متر با نقطه شروع تصادفی در منطقه مورد مطالعه قرار گرفت که در نتیجه آن، تعداد ۴۶ نقطه نمونه‌برداری برای اجرای روش k -NN انتخاب شدند.

روش پژوهش

روش k -NN به عنوان یکی از مطلوب‌ترین روش‌های فاصله‌ای در آماربرداری جنگل با استفاده از اندازه‌گیری فاصله تا k امین نزدیکترین درخت از n نقطه نمونه‌برداری، تراکم را برآورد می‌کند (Magnussen *et al.*, 2012a; Seo *et al.*, 2014). در مطالعات پیشین از تعداد k درخت مختلف در برآورد تراکم درختان استفاده شده است (جدول ۱). در این پژوهش با توجه به متوسط فاصله درختان از یکدیگر (۱۰/۷ متر) و همچنین ابعاد شبکه آماربرداری (100×100 متر)، تعداد k برابر با دو تا ۱۰ انتخاب شد تا پس از انتخاب درخت دهم، بین نقاط نمونه‌برداری تداخل بوجود نیاید.

برآوردکننده‌های ارائه شده در پیشینه پژوهش، نسبت به الگوی مکانی درختان در جنگل از حساسیت بیشتری برخوردارند و کارایی آنها تأیید شده است. نتایج این پژوهش امکان استفاده از روش k -NN بهینه را در مطالعات آینده فراهم خواهد ساخت.

مواد و روش‌ها

برای انجام این مطالعه، جنگل تحقیقاتی بنه فیروزآباد در استان فارس انتخاب شد. این مجموعه که از سال ۱۳۷۵ به عنوان ایستگاه تحقیقاتی بنه شناخته شده است، مساحتی بالغ بر ۹۳۷۴ هکتار دارد و در جنوب غربی استان فارس و در نزدیکی شهرستان فیروزآباد واقع شده است (Owji & Hamzepour, 2012). موقعیت مکانی این عرصه در طول جغرافیائی $30^{\circ} 52'$ تا $40^{\circ} 52'$ شرقی و عرض جغرافیائی 29° تا $15' 29^{\circ}$ شمالی قرار دارد. برای اجرای پژوهش پیش‌رو، بخشی از این جنگل تحقیقاتی با وسعت ۴۵ هکتار که بنه تنها گونه درختی آن است، انتخاب شد که از جنوب به جاده شیراز-کازرون، از شرق و غرب به اراضی کشاورزی و از شمال به کوه منتهی می‌شود. پس از

جدول ۱- تعداد k مورد استفاده و تراکم منطقه مورد مطالعه در برخی از پژوهش‌های پیشین

منبع	تعداد k	تراکم (در هکتار)	منبع	تعداد k	تراکم (در هکتار)
Jahan و همکاران (۲۰۱۰)	۹	۱۹۴	Kleinn و Vilcko (۲۰۰۶)	۱ تا ۱۳	۱۰۰
Karamshahi و همکاران (۲۰۱۲)	۵	۱۴۰	Magnussen و همکاران (۲۰۱۲a)	۳ تا ۸	از ۱۰۰ تا ۸۰۰

نزدیکترین فرد (NI)

اولین راهکار انتخاب درختان در روش k -NN، راهکار NI است که در آن، k امین نزدیکترین درخت به نقطه نمونه-برداری در نظر گرفته می‌شود (شکل ۱- الف)، سپس فاصله مستقیم هر درخت تا نقطه نمونه‌برداری اندازه‌گیری شده (r) و میانگین آنها در برآورد تراکم با استفاده از رابطه ۱ به کار می‌رود (Magnussen *et al.*, 2012a; Kleinn & Vilcko, 2006).

(2006).

$$\lambda_k = (k - 1) \times \pi^{-1-0.5} \sqrt{\bar{r}_k \bar{r}_{k-1}} \quad (1) \text{ رابطه (۱)}$$

در رابطه فوق: λ_k تراکم درختان براساس k امین درخت، \bar{r}_k میانگین فاصله تا k امین درخت و \bar{r}_{k-1} میانگین فاصله تا $k-1$ امین درخت است. نزدیکترین همسایه (NN)

روش NN دومین راهکار انتخاب درختان در روش k -

پس از تقسیم نقطه نمونه برداری در چهار جهت اصلی، k امین نزدیکترین درخت به نقطه نمونه برداری در هر چارک انتخاب شده و فاصله تا آن اندازه گیری می شود (شکل ۱-د)، سپس با استفاده از یک برآوردکننده ناریب در راهکار PCQ (رابطه ۲)، تراکم برآورد می شود (Hijbeek *et al.*, 2013).

$$\lambda_k = 4(4n - 1) / \pi \sum_{i=1}^n \sum_{j=1}^4 r_{ij}^2 \quad \text{رابطه (۲)}$$

در رابطه فوق: λ_k تراکم درختان براساس k امین درخت، r_{ij} فاصله تا درخت j در نقطه نمونه برداری i و n تعداد نقاط نمونه برداری است.

همسایه چارکی (QN)

راهکار QN که برای نخستین بار توسط Zhang و Zhu (۲۰۰۹) ارائه شد، به راهکار PCQ شبیه است، با این تفاوت که k امین نزدیکترین همسایه از نزدیکترین درخت به نقطه نمونه برداری مبنای برآورد تراکم است (شکل ۱-ه). برای برآورد تراکم با استفاده از راهکار QN از رابطه ۲ استفاده می شود.

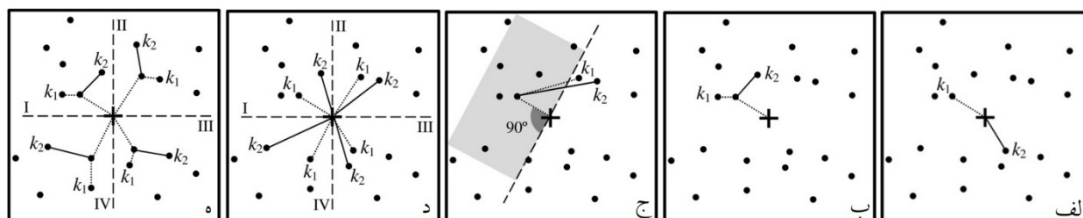
NN است که در آن k امین نزدیکترین درخت همسایه به نزدیکترین درخت به نقطه نمونه برداری در نظر گرفته می شود (شکل ۱-ب). فاصله تا نزدیکترین همسایه مبنای برآورد تراکم است که میانگین آن در رابطه ۱ که در راهکار NI به کار رفت، قرار داده می شود. رابطه ۱ یکی از بهترین برآوردکننده های ناریب در روش k -NN است که در پژوهش های پیشین مورد تأیید قرار گرفته است (Hill *et al.*, 2005).

جفت های تصادفی (RP)

در راهکار RP نزدیکترین درخت به نقطه نمونه برداری متصل شده و خطی عمود بر آن برازش داده می شود. سپس k امین نزدیکترین همسایه در طرف مقابل خط انتخاب می شود (شکل ۱-ج). میانگین فاصله های اندازه گیری شده در رابطه ۱ به کار رفته و تراکم برآورد می شود (Hill *et al.*, 2005).

چارک نقطه مرکز (PCQ)

راهکار PCQ یکی از پرکاربردترین راهکارها در برآورد ویژگی های ساختاری جنگل مانند تراکم است (Bryant *et al.*, 2004; Guebas & Koedam, 2006). در این راهکار



شکل ۱- راهکارهای نزدیکترین فرد (NI) (الف)، نزدیکترین همسایه (NN) (ب)، جفت های تصادفی (RP) (ج)، چارک نقطه مرکز (PCQ) (د) و

همسایه چارکی (QN) (ه) در روش k -NN (شکل از نگارندگان)

استفاده شد که هر کدام در مطالعات پیشین مورد استفاده و تأیید قرار گرفته است (Fehrmann & Klein, 2005; Baffetta *et al.*, 2011; Magnussen *et al.*, 2012a).

$$\text{RMSE} = \sqrt{(\sum (\lambda_k - \lambda_r)^2) / n} \quad \text{رابطه (۳)}$$

$$\text{Bias} = \lambda_k - \lambda_r \quad \text{رابطه (۴)}$$

در رابطه های ۳ و ۴، λ_k تراکم درختان براساس k امین

بررسی صحت و دقت نتایج

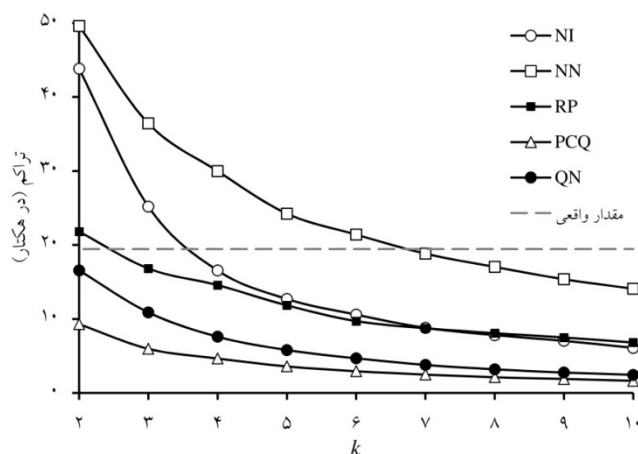
به منظور ارزیابی دقت نتایج به دست آمده از کاربرد روش k -NN با k ها و راهکارهای مختلف در برآورد تراکم درختان بنه در منطقه مورد مطالعه، از دو معیار مهم خطای یعنی مجذور میانگین مربعات خطا (Root mean squared error: RMSE) (رابطه ۳) و اریبی (Bias) (رابطه ۴)

$$z = \frac{r_{\text{obs}} - r_{\text{exp}}}{SEr_{\text{exp}}} \quad \text{رابطه (۶)}$$

در رابطه فوق: مخرج کسر انحراف معیار فاصله‌های مشاهده شده در حالت تصادفی است. اگر مقدار z بین $\pm 1/96$ باشد، اختلاف معنی‌داری بین الگوی مکانی مشاهده شده و حالت تصادفی وجود ندارد. در حالی که بیشتر بودن z از $1/96+$ نشان دهنده الگوی مکانی پراکنده و کمتر بودن آن از $1/96-$ نشان دهنده الگوی مکانی کپه‌ای است.

نتایج

آماربرداری صد درصد نشان داد که ۸۷۵ درخت بنه دانه‌زاد با قطر بیشتر از ۵ سانتی‌متر در منطقه مورد مطالعه وجود دارد که با توجه به مساحت آن، تراکم منطقه مذکور $19/44$ درخت بنه در هکتار با توزیع کاملاً کپه‌ای (مقدار شاخص نزدیکترین همسایه $0/79$ و مقدار z برابر با $-12/38$) می‌باشد. سپس از ۴۶ نقطه نمونه‌برداری، فاصله تا k امین نزدیکترین درخت اندازه‌گیری شد و تراکم درختان بنه برآورد شد که نتایج آن در شکل ۲ ارائه شده است. نتایج نشان داد که به‌طور کلی با افزایش تعداد k ، تراکم برآورد شده در همه راهکارها کاهش می‌یابد و با افزایش k در چهار راهکار NI، RP، PCQ و QN تراکم برآورد شده کمتر از ۱۰ درخت در هکتار شده است که از مقدار واقعی فاصله داشتند. همچنین در مقدار k زیاد، نتایج دو راهکار NI با RP و PCQ با QN بسیار به هم نزدیک شدند (شکل ۲).



شکل ۲- تراکم درختان بنه برآورد شده به‌وسیله روش k -NN با k ها و راهکارهای مختلف

درخت، λ تراکم واقعی و n تعداد نقاط نمونه‌برداری است. همچنین برای بررسی صحت نتایج از آزمون t تک‌نمونه‌ای در سطح اطمینان ۹۵ درصد استفاده شد.

الگوی مکانی درختان بنه

از آنجایی که روش‌های فاصله‌ای از جمله k -NN به نحوه توزیع درختان در جنگل حساس هستند و در الگوهای مکانی مختلف، عملکرد متفاوتی دارند (Magnussen, 2012b; Magnussen et al., 2012)، آگاهی از پراکنش درختان بنه در منطقه مورد مطالعه ضروری به نظر می‌رسد. بنابراین از شاخص نزدیکترین همسایه (رابطه ۵) استفاده شد که علاوه بر تأیید آن در مطالعات پیشین (Erfanifard & Mahdian, 2012)، الگوی مکانی را براساس فاصله تا نزدیکترین همسایه تعیین کرده که از نظر ساختاری با روش k -NN مشابه است.

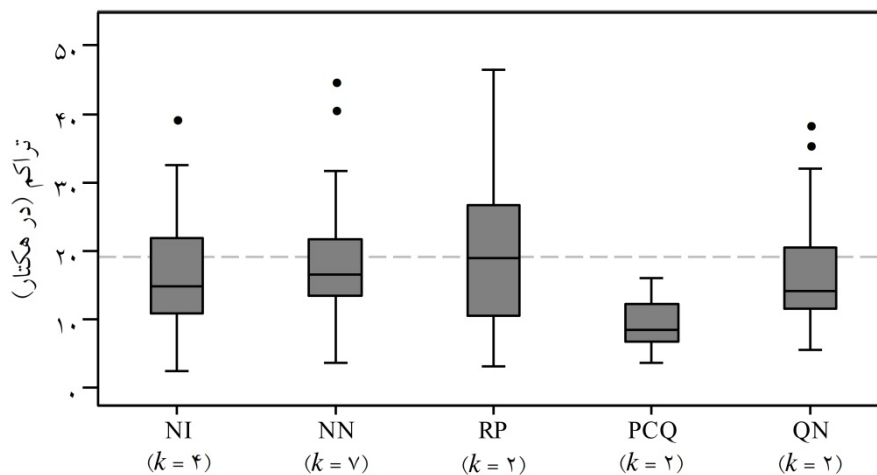
$$R = \frac{r_{\text{obs}}}{r_{\text{exp}}} \quad \text{رابطه (۵)}$$

در رابطه فوق: r_{obs} فاصله تا نزدیکترین همسایه مشاهده شده و r_{exp} فاصله تا نزدیکترین همسایه در حالت تصادفی است. در صورتی که مقدار شاخص برابر با یک باشد، الگوی مکانی درختان تصادفی است و در الگوهای کپه‌ای و پراکنده، مقدار شاخص به ترتیب کمتر و بیشتر از یک است. نتایج با آزمون آماری z (رابطه ۶) نیز ارزیابی می‌شود (Diggle, 2003).

بود. برخلاف راهکارهای RP و QN، توزیع تراکم‌های برآورد شده با k های مطلوب در راهکارهای NI ($k=4$)، NN ($k=7$) و PCQ ($k=2$) تقریباً شبیه توزیع نرمال بود. کمترین میزان پراکنندگی تراکم برآوردی مربوط به راهکار PCQ و بیشترین آن مربوط به راهکار RP بود. نتایج مقایسه محدوده تراکم‌های برآورد شده در k های مطلوب در هر راهکار نشان داد که به جز روش PCQ، مقدار واقعی تراکم در چارک سوم تراکم‌های برآورد شده در هر راهکار بوده است. بررسی دقت تراکم‌های برآورد شده به وسیله روش k -NN با k ها و راهکارهای مختلف نشان داد که مقدار خطا (RMSE و اریبی) $k=4$ در راهکار NI کمتر از سایر k ها بوده است (شکل ۴- الف). هرچند مقدار RMSE در راهکارهای NN و RP به ترتیب در k های نه و پنج حداقل بود، اما اریبی k های هفت و دو حداقل بوده است (شکل‌های ۴- ب و ۴- ج). روند تغییرات RMSE و اریبی در راهکارهای PCQ و QN شبیه به یکدیگر بود و در $k=2$ خطای هر دو راهکار کمتر از سایر k ها بود (شکل‌های ۴- د و ۴- ه).

نتایج ارزیابی صحت تراکم برآورد شده به وسیله روش k -NN با k ها و راهکارهای مختلف با استفاده از آزمون آماری اختلاف مشاهده شده بین مقدار واقعی و برآوردی نشان داد که در راهکار NI، تراکم برآورد شده در $k=4$ (۱۷/۵۵ درخت در هکتار) با مقدار واقعی اختلاف معنی‌داری نداشته است. همچنین راهکارهای NN و RP به ترتیب در k های هفت (۱۸/۸۲ درخت در هکتار) و دو (۲۰/۸۱) تراکم را بدون اختلاف معنی‌دار با مقدار واقعی برآورد کردند، در حالی که راهکار PCQ نتوانست به طور صحیح تراکم را در k های متفاوت برآورد کند، اگرچه اختلاف تراکم برآورد شده با راهکار QN در $k=2$ (۱۷/۵۶ درخت در هکتار) با مقدار واقعی اختلاف معنی‌داری نداشت (جدول ۲).

شکل ۳ نشان داد که تراکم‌های برآورد شده با $k=2$ در راهکار RP، تقریباً متقارن هستند و میانه آنها در وسط داده‌ها قرار دارد. در بقیه راهکارهای روش k -NN با k مطلوب تراکم‌های برآورد شده به چپ چولگی داشتند که بیانگر فراوانی بیشتر تراکم‌های کم نسبت به تراکم‌های زیاد

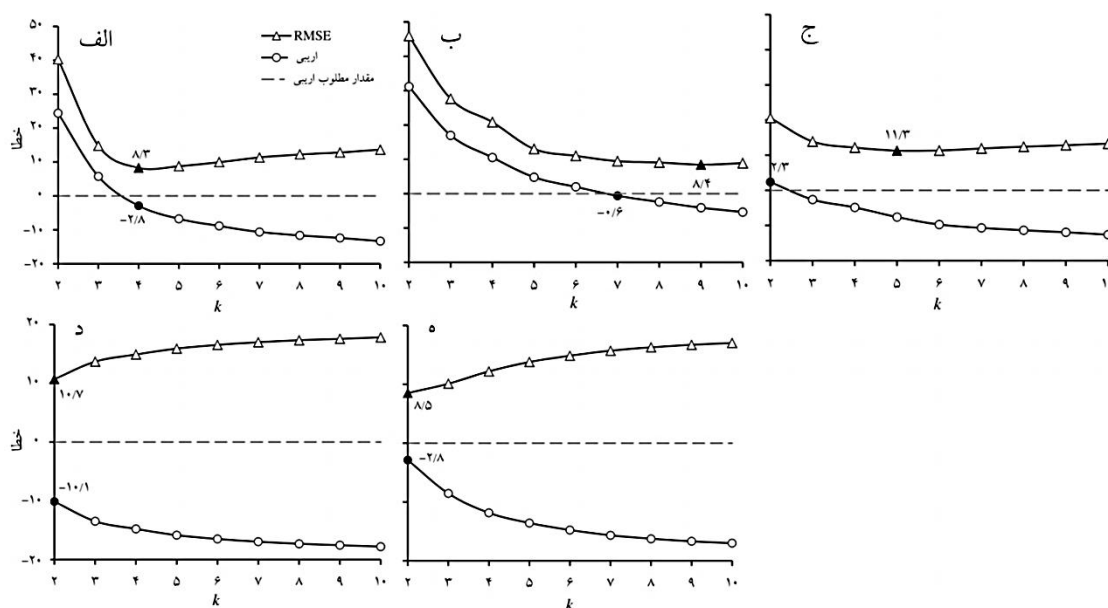


شکل ۳- نمودار جعبه‌ای تراکم برآورد شده به وسیله روش k -NN با k مطلوب در هر راهکار (خط چین خاکستری مقدار واقعی تراکم را نشان می‌دهد)

جدول ۲- نتایج ارزیابی صحت تراکم (در هکتار) برآورد شده به وسیله روش k -NN با k ها و راهکارهای مختلف

	NI تراکم (معنی‌داری)	NN تراکم (معنی‌داری)	RP تراکم (معنی‌داری)	PCQ تراکم (معنی‌داری)	QN تراکم (معنی‌داری)
$k=2$	۴۳/۸۳* (۰/۰۰۰)	۵۰/۶۱* (۰/۰۰۰)	۲۰/۸۱ ^{ns} (۰/۴۵۳)	۹/۳۰* (۰/۰۰۰)	۱۷/۵۶ ^{ns} (۰/۰۶۲)
$k=3$	۲۵/۱۹* (۰/۰۰۷)	۳۶/۴۳* (۰/۰۰۰)	۱۶/۸۱* (۰/۰۲۱)	۵/۹۶* (۰/۰۰۰)	۱۰/۸۹* (۰/۰۰۰)
$k=4$	۱۷/۵۵ ^{ns} (۰/۰۶۱)	۲۹/۹۸* (۰/۰۰۰)	۱۴/۵۷* (۰/۰۰۷)	۴/۶۵* (۰/۰۰۰)	۷/۶۰* (۰/۰۰۰)
$k=5$	۱۲/۷۰* (۰/۰۰۰)	۲۴/۲۳* (۰/۰۱۱)	۱۱/۸۳* (۰/۰۰۰)	۳/۵۹* (۰/۰۰۰)	۵/۸۷* (۰/۰۰۰)
$k=6$	۱۰/۵۸* (۰/۰۰۰)	۲۱/۴۰* (۰/۰۲۳)	۹/۷۱* (۰/۰۰۰)	۲/۹۴* (۰/۰۰۰)	۴/۶۸* (۰/۰۰۰)
$k=7$	۸/۸۰* (۰/۰۰۰)	۱۸/۸۳ ^{ns} (۰/۶۶۲)	۸/۷۵* (۰/۰۰۰)	۲/۴۸* (۰/۰۰۰)	۳/۷۹* (۰/۰۰۰)
$k=8$	۷/۸۰* (۰/۰۰۰)	۱۷/۰۴* (۰/۰۰۷)	۸/۰۸* (۰/۰۰۰)	۲/۱۲* (۰/۰۰۰)	۳/۱۹* (۰/۰۰۰)
$k=9$	۷/۰۴* (۰/۰۰۰)	۱۵/۳۶* (۰/۰۰۱)	۷/۴۹* (۰/۰۰۰)	۱/۸۸* (۰/۰۰۰)	۲/۷۶* (۰/۰۰۰)
$k=10$	۶/۱۰* (۰/۰۰۰)	۱۴/۰۸* (۰/۰۰۰)	۶/۸۲* (۰/۰۰۰)	۱/۶۴* (۰/۰۰۰)	۲/۴۴* (۰/۰۰۰)

* معنی‌دار در سطح اطمینان ۹۵ درصد؛ ^{ns} غیرمعنی‌دار



شکل ۴- تغییرات خطا (RMSE و اریبی) در روش k -NN با راهکارهای NI (الف)، NN (ب)، RP (ج)، PCQ (د) و QN (ه)

بحث

تراکم (تعداد در هکتار) یکی از مهمترین ویژگی‌های کمی هر جامعه جنگلی است که برآورد آن، هدف انواع روش‌های نمونه‌برداری به‌ویژه روش‌های فاصله‌ای بوده است (Hijbeek *et al.*, 2013) و کارایی آنها در جنگل‌های زاگرس نیز تأیید شده است (Askari *et al.*, 2013). یکی از روش‌های مهم فاصله‌ای که در اندازه‌گیری تراکم در جنگل کاربرد فراوانی یافته است، روش k -NN است که به دلیل قابلیت برآورد تراکم در بوم‌سازگان‌های مختلف و همچنین حداقل تأثیرپذیری از الگوی پراکنش درختان در جنگل یافته است و کارایی آن در جنگل‌های زاگرس نیز مورد تأیید قرار گرفته است (Jahan *et al.*, 2010; Karamshahi *et al.*, 2012). با وجود مطالعات پیشین در مورد کاربرد روش k -NN برای سنجش ویژگی‌های ساختاری جنگل از جمله تراکم، راهکارهای انتخاب درخت k جز در پژوهش‌های معدود (Kleinn & Vilcko, 2006) کمتر مورد توجه قرار گرفته است. بنابراین مقایسه روش k -NN با راهکارهای مختلف و همچنین بررسی تأثیر آنها بر کارایی این روش ضروری به نظر می‌رسد. پژوهش پیش‌رو با هدف معرفی و کاربرد انواع راهکارهای انتخاب k امین نزدیکترین همسایه در روش k -NN به منظور برآورد تراکم درختان بنه در بخشی از جنگل‌های زاگرس در استان فارس انجام شد. همچنین تأثیر این راهکارها بر صحت و دقت تراکم برآورد شده به وسیله روش k -NN با k های متفاوت مطالعه شد.

پژوهش پیش‌رو در یک توده تنک به مساحت ۴۵ هکتار از جنگل تحقیقاتی بنه فیروزآباد با تراکم ۱۹/۴۴ درخت بنه در هکتار و توزیع کاملاً کپه‌ای انجام شد. تعداد کم درختان در جامعه جنگلی مورد مطالعه بر کارایی روش‌های فاصله‌ای از جمله روش k -NN تأثیر می‌گذارد (Kleinn & Vilcko, 2006; Magnussen *et al.*, 2012a). چنانچه در جدول ۱ ملاحظه شد، کارایی روش k -NN در برخی از مطالعات پیشین که به‌عنوان مثال آورده شده‌اند، در تراکم‌های بیشتر از ۱۰۰ درخت در هکتار تأیید شده است،

در حالی‌که نتایج پژوهش پیش‌رو کارایی روش مذکور را در تراکم‌های بسیار کمتر (حدود ۲۰ درخت در هکتار) نیز تأیید کرد. همچنین این پژوهش نیز مانند مطالعات پیشین (Magnussen *et al.*, 2012a) قابلیت روش k -NN را در برآورد تراکم یک منطقه با آرایش کپه‌ای درختان تأیید کرد.

در مطالعات پیشین راهکارهای NI (Engeman *et al.*, 1994)، NN (Temesgen *et al.*, 2008; McRoberts, 2012)، RP (Hill *et al.*, 2005)، PCQ (Hijbeek *et al.*, 2004)، QN (Zhu & Zhang, 2013; Bryant *et al.*, 2011) و Leavitt & Clair, 2009) مورد استفاده قرار گرفته است، اما به مقایسه این راهکارها و تأثیر آنها بر عملکرد روش k -NN پرداخته نشده است. اگرچه در پژوهش Kleinn و Vilcko (۲۰۰۵) به مقایسه روش k -NN با دو راهکار NI و PCQ پرداخته شده است و هر دو تأیید شده‌اند، در پژوهش پیش‌رو روش مذکور با راهکار PCQ نتوانست تراکم را به‌درستی برآورد کند (جدول ۲). دلیل این موضوع ممکن است به عدم حساسیت راهکار مذکور به الگوی مکانی درختان و ناتوانی برآورد تراکم درختان با الگوی مکانی کپه‌ای و همچنین ناکارآمدی در توده‌های با تراکم کم درختان مربوط باشد. نتایج پژوهش پیش‌رو نشان داد که روش k -NN با راهکارهای مورد بررسی از عملکرد متفاوتی در شرایط یکسان (توده درختان بنه با تراکم و شبکه نمونه‌برداری مشابه و آرایش کپه‌ای) برخوردار بوده است (جدول ۲). همچنین نتایج نشان داد که صحت روش k -NN با هر راهکار در k های متفاوت تأیید شده است (جدول ۲ و شکل ۳) که این نتیجه‌گیری با مطالعات پیشین (Kleinn & Vilcko, 2006; Magnussen *et al.*, 2012a) مطابقت دارد و برخلاف برخی مطالعات، نمی‌توان از یک k مشخص در روش k -NN استفاده کرد. لازم به توضیح است که در پژوهش پیش‌رو فاصله تا k مین نزدیکترین همسایه از k برابر یک تا ۱۰ اندازه‌گیری شد، اما از آنجایی‌که در رابطه ۱ نمی‌توان تراکم را برای $k=1$ تعیین کرد، نتایج از $k=2$ ارائه شده است.

به‌منظور برآورد تراکم با استفاده از روش k -NN، برآورد

کننده‌های زیادی طراحی و به‌کار برده شده‌اند. استفاده از برآوردکننده مناسب و نارایب نیز بر کارایی روش k -NN و صحت نتایج آن تأثیر فراوانی دارد. در پژوهش پیش‌رو برای پنج راهکار مورد بررسی از دو برآوردکننده‌ای استفاده شد که براساس مطالعات پیشین، در انواع الگوهای مکانی از صحت و دقت قابل قبولی برخوردار بودند و ویژگی‌های ساختاری جامعه‌های مورد بررسی را برآورد کردند (Kleinn & Vilcko, 2006; Magnussen *et al.*, 2012a;) (Hijbeek *et al.*, 2013).

سیاسگزاری

مطالعات میدانی این پژوهش با مجوز مرکز تحقیقات کشاورزی و منابع طبیعی استان فارس انجام شده است. علاوه بر این، از حمایت مالی معاونت آموزشی و تحصیلات تکمیلی دانشگاه شیراز (کد اعتبار ۹۰GCU۲M۱۵۳۱۲۶) در انجام این پژوهش استفاده شد که بدین‌وسیله سپاسگزاری می‌شود.

References

- Askari, Y., Zobeiri, M. and Sohrabi, H., 2013. Comparison of five distance sampling methods for estimating quantitative characteristics of Zagros forests. *Iranian Journal of Forest and Poplar Research*, 21(2): 316-328 (In Persian).
- Baffetta, F., Corona, P. and Fattorini, L., 2011. Design-based diagnostics for k -NN estimators of forest resources. *Canadian Journal of Forest Resources*, 41: 59-72.
- Bonham, C.D., 2013. *Measurements for terrestrial vegetation*. Wiley-Blackwell Publishers, West Sussex, 246p.
- Bryant, D.M., Ducey, M.J., Innes, J.C., Lee, T.D., Eckert, R.T. and Zarin, D.J., 2004. Forest community analysis and the point-centered quarter method. *Plant Ecology*, 175: 193-203.
- Chirici, G., Winter, S. and McRoberts, R.E., 2011. *National Forest Inventories: Contributions to Forest Biodiversity Assessments*. Springer, New York, 226p.
- Diggle, P.J., 2003. *Statistical analysis of spatial point patterns*. Arnold Publishers, New York, 159p.
- Engeman, R.M., Sugihara, R.T., Pank, L.F. and Dusenberry, W.E., 1994. A comparison of plotless density estimators using Monte Carlo simulation. *Ecology*, 75(6): 1769-1779.
- Erfanifard, Y. and Mahdian, F., 2012. Comparative investigation on the methods of true spatial pattern analysis of trees in forests, Case study: Wild pistachio research forests,

علاوه بر ارزیابی صحت تراکم برآورد شده در این مطالعه که به دلیل دسترسی به تراکم واقعی امکان‌پذیر شد (برخلاف برخی از مطالعات پیشین) و مقایسه کارایی راهکارهای مورد مطالعه، دقت آنها نیز با دو معیار متداول خطا (RMSE و اریبی) بررسی شد. اگرچه دو معیار خطا در راهکارهای NI، PCQ و QN در k های مشابه حداقل بودند، اما در دو راهکار NN و RP نتایج متفاوتی داشتند (شکل ۴). در این پژوهش در راهکارهای NN و RP معیار اریبی برای انتخاب k مطلوب در هر راهکار مورد توجه قرار گرفت (شکل ۴)، زیرا با نتایج ارزیابی صحت با آزمون آماری ($k=7$ در NN و $k=2$ در RP) مطابقت داشتند (جدول ۲). این موضوع نشان داد که علاوه بر معیارهای خطا در مطالعه دقت نتایج به‌دست‌آمده از روش k -NN باید از بررسی صحت با آزمون‌های مناسب آماری نیز استفاده شود و سپس به نتیجه‌گیری نهایی اقدام گردد.

به‌طور کلی نتایج پژوهش پیش‌رو نشان داد که در صورت استفاده از k و راهکار مناسب، روش k -NN با صحت و دقت قابل قبول قادر به برآورد تراکم درختان بنه با الگوی مکانی کپه‌ای در یک توده تنک (۱۹/۴۴) درخت در هکتار) در جنگل‌های زاگرس بوده است. یافته‌های این مطالعه تأیید کرد که مقدار k و شکل انتخاب k امین نزدیکترین همسایه بر کارایی روش k -NN تأثیر گذاشته است. همچنین نتایج نشان داد که استفاده همزمان از دقت و صحت در انتخاب k و راهکار بهینه روش k -NN ضروری به‌نظر می‌رسد. اگرچه نتایج پژوهش پیش‌رو بیانگر آن بود

- Kung, Y.H., Lin, P.S. and Kao, C.H. 2012. An optimal k -nearest neighbor for density estimation. *Statistics and Probability Letters*, 82: 1786-1791.
- Leavitt, S.D. and Clair, L.L., 2011. Estimating *Xanthoparmelia* (Parmeliaceae) population density in subalpine communities in southern Utah, U.S.A. using two distance methods, with implications for assessing community composition. *Bryologist*, 114(3): 625-636.
- Magnussen, S., 2012. A new composite k -tree estimator of stem density. *European Journal of Forest Research*, 131: 1513-1527.
- Magnussen, S., Fehrman, L. and Platt, W.J., 2012a. An adaptive composite density estimator for k -tree sampling. *European Journal of Forest Research*, 131: 307-320.
- Magnussen, S., McRoberts, R.E. and Tomppo, E.O., 2012b. A resampling variance estimator for the k nearest neighbours technique. *Canadian Journal of Forest Research*, 40: 648-658.
- McRoberts, R.E., 2012. Estimating forest attribute parameters for small areas using nearest neighbors techniques. *Forest Ecology and Management*, 272: 3-12.
- Owji, M.G. and Hamzepour, M., 2012. Vegetation profile of wild pistachio experimental forest. Published by Research Institute of Forests and Rangelands, Tehran, 240p (In Persian).
- Temesgen, H., Barrett, T.M. and Latta, G., 2008. Estimating cavity tree abundance using Nearest Neighbor imputation methods for western Oregon and Washington forests. *Silva Fennica*, 42(3): 337-354.
- Pourreza, M., Shaw, J.D. and Zangeneh, H., 2008. Sustainability of wild pistachio (*Pistacia atlantica* Desf.) in Zagros forests, Iran. *Forest Ecology and Management*, 255: 3667-3671.
- Seo, H.S., Phua, M.H., Ong, R., Choi, B. and Lee, J.S., 2014. Determining aboveground biomass of a forest reserve in Malaysian Borneo using k -Nearest Neighbour method. *Journal of Tropical Forest Science*, 26(1): 58-68.
- Zhu, X. and Zhang, J., 2009. Quartered neighbor method: a new distance method for density estimation. *Frontiers of Biology in China*, 4(4): 574-578.
- Fars province, Iran. *Iranian Journal of Forest and Poplar Research*, 20(1): 62-73 (In Persian).
- Fehrmann, L. and Kleinn, C., 2005. A k -Nearest Neighbour approach for estimating of single-tree biomass. *Proceedings of the 7th FIA Symposium, USA, 3-6 Oct. 2005*: 251-259.
- Guebas, F.D. and Koedam, N., 2006. Empirical estimate of the reliability of the use of the Point-Centered Quarter Method (PCQM): Solutions to ambiguous field situations and description of the PCQM+ protocol. *Forest Ecology and Management*, 228: 1-18.
- Hanberry, B.B., Yang, J., Kabrick, J.M. and He, H.S., 2012. Adjusting forest density estimates for surveyor bias in historical tree surveys. *American Midland Naturalist*, 167: 285-306.
- Hijbeek, R., Koedam, N., Khan, M.N.I., Kairo, J.G., Schoukens, J. and Guebas, F.D., 2013. An evaluation of plotless sampling using vegetation simulations and field data from mangrove forest. *PLoS ONE*, 8(6): 1-10.
- Hill, D., Fasham, M., Tucker, G., Shewry, M. and Shaw, P., 2005. *Handbook of Biodiversity Methods*. Cambridge University Press, Cambridge, 573p.
- Jahan, A., Akhavan, R. and Erfanifard, Y., 2010. Application of k -NN (9-tree) sampling method in central Zagros forests (Case study: Servak forest, Yasouj). *The 1st Symposium on Threats and Destroying Factors of Biodiversity in Central Zagros. Iran, 16-17 Feb. 2010*: 231-233 (In Persian).
- Jazirehi, M.H. and Ebrahimi Rostaghi, M., 2003. *Silviculture in Zagros*. University of Tehran Press, Tehran, 560p (In Persian).
- Kangas, A. and Maltamo, M., 2006. *Forest Inventory: Methodology and Applications*. Springer, 365p.
- Karamshahi, A., Zobeiri, M., Namiranian, M. and Fegghi, J., 2012. Investigation on application of k -NN (k -Nearest Neighbor) sampling method in Zagros forests (Case study: Karzan forest, Ilam). *Iranian Journal of Forest and Poplar Research*, 19(4): 453-46 (In Persian).
- Kleinn, C. and Vilcko, F. 2005. On estimation in k -tree sampling. *Proceedings of the 7th FIA Symposium. USA, 3-6 Oct. 2005*: 203-208.
- Kleinn, C. and Vilcko, F., 2006. A new empirical approach for estimation in k -tree sampling. *Forest Ecology and Management*, 237: 522-533.

Comparing different k -NN sampling methods for density estimation of wild pistachio (*Pistacia atlantica* Desf.) with clustered spatial pattern in a Zagros open stand

M. Moselou¹ and Y. Erfanifard^{2*}

1- M.Sc. Student, College of Agriculture, Shiraz University, Shiraz, Iran

2* - Corresponding author, Assistant Prof., College of Agriculture, Shiraz University, Shiraz, Iran

E-mail: erfanifard@shirazu.ac.ir

Received: 25.10.2014

Accepted: 20.04.2015

Abstract

Density (i.e. number of trees per unit area) is one of the important structural attributes of forest stands which is appropriate to understand forest dynamics. The k -Nearest Neighbour (k -NN) method is a distance sampling method which is commonly used to estimate quantitative attributes in forest inventories. In this study, the k -NN method with five strategies of Nearest Individual (NI), Nearest Neighbour (NN), Random Pairs (RP), Point-Centered Quarter (PCQ), and Quartered Neighbour (QN) was used to estimate the density of wild pistachio (*Pistacia atlantica* Desf.) in Zagros woodlands. A natural stand of 45 ha was selected in Bane Research Site, and was fully callipered to derive the true density. The spatial distribution of trees was clustered ($R=0.79$ and $z = -12.38$) with a true density of 19.44 trees per ha. While applying the k -NN method, different strategies as well as k ranging between 2 and 10 were tested across 46 sample points in a $100 \times 100 m^2$ sampling grid. The results showed that all strategies except PCQ significantly estimated the density at $\alpha=0.05$. Furthermore, the number of k and the strategy of k -tree selection affected the accuracy and precision of k -NN results, since NI in $k=4$, NN in $k=7$, RP and QN in $k=2$ estimated the density with the least error (RMSE and Bias). In conclusion, the optimum k -NN method with appropriate k and strategy could estimate the density of wild pistachio trees with clustered spatial distribution in an open stand in Zagros woodlands.

Keywords: Wild pistachio, density, Zagros, error, k -NN.