ISSN (1681- 6870)

تصنيف الاراضي الزراعية حسب المساحة وكمية الانتاج لبعض محاصيل الخضر في محافظات العراق لعامي (2005) و(2010) بأستخدام طرائق التحليل العنقودي (Ward's والعنقدة المركزية الهرمية)

محمد عبد الودود محمد
mohammedabdulwadood1@gmail.com
أ.م.د اسماء غالب الراوي
asmaa_alhasen@yahoo.com
جامعة بغداد - كلية الادارة والاقتصاد - قسم الإحصاء - بغداد - العراق

المستخلص

يعد القطاع الزراعي من القطاعات الاقتصادية المهمة التي تسهم وبدور كبير في تحقق التنمية الاقتصادية في البلدان النامية اذ يسهم في الجزء الأكبر من الناتج المحلي الإجمالي لهذه البلدان لذا ولأهمية هذا القطاع الحيوي سعى هذا البحث الى تصنيف الأراضي الزراعية غير المصنفة سابقا للمحافظات كافة عدا محافظات إقليم كردستان لمدى صلاحيتها لزراعة الخضر لعامي (2005) و(2010)، اذ يعدان عامين مهمين جدا للقطاع الزراعي في العراق، اذ العام (2005) يأتي بعد الانهيار في هذا القطاع نظرا للفراغ السياسي والاقتصادي بعد عام (2003) في حين يشكل العام (2010) بداية النهوض للقطاع الزراعي من خلال وضع سياسات انتاجية وخطط خمسية من قبل وزارة الزراعة، واستخدم لحل هذه المشكلة اسلوبا (Ward's) والعنقدة المركزية الهرمية للتحليل العنقودي، وايضا يسعى البحث الى التنبؤ بكميات الانتاج ما بعد عام (2010)، وتوصل البحث الى ان المحافظات الشمالية عموما ونينوى خصوصا تعد من اخصب وافضل الاراضي الزراعية في العراق نظراً لكميات انتاجها مقارنة مع باقي المحافظات رغم التراجع الكبير في الاهتمام في القطاع الزراعي وايضا لتصدرها في تكوين العناقيد قبل المحافظات الاخرى.

المصطلحات الرئيسية للبحث: تحليل عنقودي، اساليب هرمية، اسلوب وارد للعنقدة، اسلوب العنقدة المركزية، تحليل انحدار.

Statistical Classification of Agricultural Land by Area and the Amount of Production of Some Agricultural Crops in Iraqi Governorates for the Years (2005) and (2010) by Using Cluster Analysis Methods (Ward's & Hierarchical Centroid Clustering)

Mohammed A. Mohammed
mohammedabdulwadood1@gmail.com
Dr. Asmaa G. AL-Rawi
asmaa_alhasen@yahoo.com
Baghdad University - College of Administration and Economic - Department of Statistics
Received 29/1/2019 **Accepted 12/5/2019**

Abstract: *The agricultural sector is an important economic sector that contributes significantly to economic development in developing countries with an important contribution bulk of the Gross domestic product of developing countries, Therefore; for the importance of this vital sector, this paper seek to classify agricultural lands not previously classified for all governorates, except the Kurdistan Region, to the extent of their validity for vegetable growing for the years (2005) and (2010). These two years are very important for the agricultural sector in Iraq, the first year (2005) comes after the collapse in this sector due to the political and economic vacuum after (2003), while the year (2010) marks the beginning of the advancement of the agricultural sector through the development of productive policies by the Ministry of Agriculture. The methods used to solve this problem are the hierarchical methods of cluster analysis in the analysis of data because these methods depend on the creation of clusters of the most similar variables and in our study these variables are the agricultural lands of the studied governorates. The paper also seeks to predict the*

ISSN (1681- 6870)

quantities of production after (2010). We found that the northern provinces in general and Ninewa in particular is one of the most fertile and best agricultural land in Iraq, compared to the quantities of production of other provinces, despite the decline in the interest in the agricultural sector and also to lead in the formation of clusters before other provinces.

Keywords: Cluster Analysis, Hierarchical Techniques, Ward's Clustering Technique, Centroid Clustering technique, Regression Analysis

1. المقدمة وهدف البحث

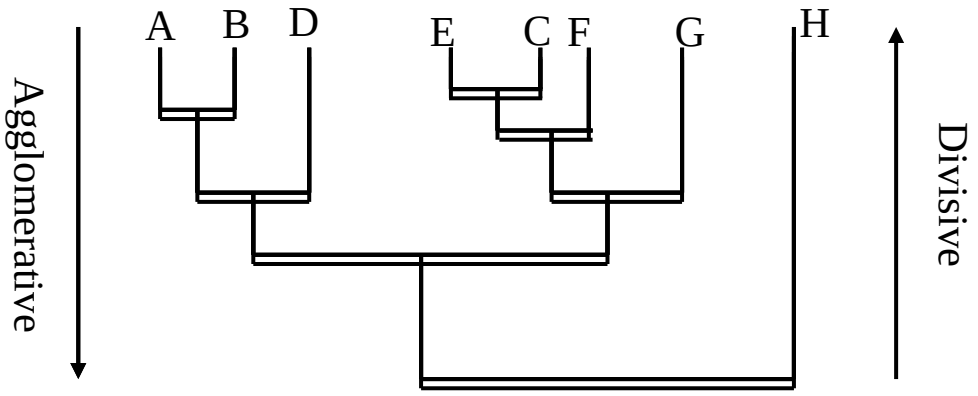
يعد العراق أحد البلدان النامية حسب تصنيف الامم المتحدة وهذه البلدان بحاجة الى موارد مالية ضخمة لغرض النهوض بمستوى الخدمات المختلفة التي تقدمها من تعليم وصحة ونقل وغيرها، ويعتبر القطاع الزراعي من أهم القطاعات الاقتصادية لكونه يوفر الغذاء للسكان ويستوعب نحو ثلث العمالة الكلية بالإضافة إلى أنه يوفر المواد الأولية الخام للصناعات العراقية نباتية كانت أم حيوانية، وايضا يعد من القطاعات الاقتصادية المهمة التي تسهم وبدور كبير في تحقق التنمية الاقتصادية في البلدان النامية ويسهم في تكوين الجزء الأكبر من الناتج المحلي الإجمالي لهذه للبلدان فضلاً عن كونه قادر على توفير القدر الأكبر من النقد الأجنبي لميزانية الدولة، وشهد هذا القطاع مثل غيره من القطاعات تراجعا كبيرا بعد عام (2003) بسبب الاهمال ونقص الدعم والتجاوزات على الاراضي وشح المياه وغيرها الكثير من المسببات لذا يهدف هذا البحث الى تصنيف الاراضي الزراعية في العراق لعامي (2005) و(2010) والتي لم تصنف سابقا للوقوف على ابرز التغيرات التي شهدتها هذه الاراضي وللوصول الى اكثر الاراضي صلاحية للزراعة في العراق من حيث كميات الانتاج وقياس آثارها وأبعادها على الواقع الزراعي في العراق ومحاولة وضع الخطط المستقبلية لتنمية وتحسين أداء هذا القطاع الحيوي والرئيسي وأخيراً محاولة إيجاد مؤشرات جديدة ذات تأثير مباشر على مثل هذه الظواهر ولم يتم أخذها بالحسبان، اذ استخدام اسلوبي (Ward's) والعنقدة المركزية الهرمية في تصنيف الاراضي الزراعية للمحافظات العراقية عدا محافظات اقليم كردستان، وأضافة الى التنبؤ بإنتاجية هذه المحاصيل للسنوات الست المقبلة إلى عام (2020) وذلك من أجل الاستفادة منها في وضع الخطط المستقبلية المناسبة لتطوير واقع هذا القطاع وتحسين أدائه، واعتمد في هذا البحث على محاصيل خضر اساسية لها تأثير كبير على المنتجين والمستهلكين المحليين وهي (الخيار، الطماطم، الباذنجان والبامية)، وتكمن أهمية البحث من كونه يدرس محاصيل الخضر الأساسية في العراق المذكورة اعلاه والمكانة البارزة لهذه المحاصيل بالنسبة للأمن الغذائي لدى المنتجين والمستهلكين في العراق، ونظراً لهذه الأهمية نلاحظ زيادة الطلب على هذه المنتجات مما أدى إلى عدم كفاية المنتج المحلي واضطرار الدولة لتعويض هذا النقص بالاستيراد من الدول المجاورة، وايضا دراسته لسنتين مؤثرتين على القطاع الزراعي في العراق فالاولى (2005) تأتي بعد الاهمال والتراجع الكبير في اداء هذا القطاع نتيجة الفراغ السياسي والاقتصادي بعد عام (2003) اما الثانية (2010) فتأتي بعد محاولة النهوض بهذا القطاع وبالتالي بداية وضع الخطط التي تساهم على تقويم اداءه، واعتمدت في التحليل بيانات الكراس الزراعي للجهاز المركزي للاحصاء لعام (2011) وقد شمل اطار هذه الدراسة محافظات العراق كافة عدا محافظات اقليم كردستان الثلاث وقد كان اطار المسح هي المدة (2002-2010) ويتمثل مجتمع الدراسة بجميع الاراضي الزراعية والحقول الزراعية الخاصة والحكومية التابعة لوزارة الزراعة، ويعد التحليل العنقودي من اهم اساليب التحليل الاحصائي متعدد المتغيرات والضرورية في مهام العنقدة والتصنيف حيث يستخدم لتصنيف المجاميع وذلك من خلال القيام بإجراءات تهدف إلى تصنيف مجموعة حالات (Cases) أو متغيرات (Variables) بطرق معينة وترتيبها داخل عناقيد (Cluster) بحيث تكون الحالات المصنفة داخل عنقود معين متجانسة فيما يتعلق بخصائص محددة وتختلف عن حالات أخرى موجودة في عنقود آخر [2]، ويتميز التحليل العنقودي بقابلية استخدامه في العديد من المجالات لذا نلاحظ ان هناك الكثير من الباحثين الذين اعتمدوا هذا الاسلوب في دراساتهم، ففي عام (1989) استخدام الباحث [جبار، 2] التحليل العنقودي للمساعدة في تشخيص مرض الـگـلازار الذي يصيب الأطفال بمرض فقر الدم وذلك عن طريق تصنيف المرضى إلى مجاميع مرضية اذ يساعد التحليل في توفير طريقة رياضية إحصائية للأطباء تقدم الكثير من الخدمات التي من شأنها أن تساعد على تشخيص أمراض عديدة أخرى، وفي عام (2002) قامت الباحثة [بيثون، 1] باستخدام التحليل العنقودي لاستخراج القيم الشاذة من خلال استخدام الصيغ والتراكيب المختلفة التي يأخذها هذا التحليل وذلك باقتراح طريقة مبسطة مقارنة مع الطرق المعتمدة حالياً في عملية الاكتشاف وكذلك مقارنة هذه الطريقة مع طرق الاكتشاف الأخرى في مجال التحليل العنقودي كالتحليل لمتغير واحد ولمتعدد المتغيرات، وفي عام (2011) قام الباحث [جوني، 3] بتصنيف مستويات المعيشة للأسر العراقية خلال الفترة (2006-2007) باستخدام التحليل العنقودي، وفي عام (2012) قامت الباحثة [عبد الحميد، 4] باستخدام التحليل العنقودي لتقدير أنموذج خطي للمرونات الانفاقية لبعض المجاميع السلعية في العراق من خلال البيانات المتاحة في مسح شبكة معرفة العراق، وفي العام نفسه قامت الباحثة [ناصر، 5] ببيان تغير الإنفاق على السلع والخدمات الرئيسية المصنفة حسب نظام الحسابات القومية خلال الفترة (1971-2010) وذلك باستعمال التحليل العنقودي وحسب المستويات البيئية لكل سنة (حضر، ريف).

2. اساليب التحليل العنقودي

للتحليل العنقودي اساليب عدة يمكن ان تصنف الى نوعين اساسيين هما عنقدة هرمية (Hierarchical Clustering) وعنقدة غيرهرمية (Non-Hierarchical Clustering) [1]، واعتمدت اساليب النوع الاول في هذا البحث، أسلوب العنقدة

ISSN (1681- 6870)

الهرمية لا يهدف إلى تقسيم البيانات إلى عدد معين من العناقيد في خطوة واحدة بل هو عبارة عن سلسلة هرمية من العناقيد المتشابكة التي تخرج من عنقود واحد يحتوي على جميع العناصر إلى (n) من العناقيد التي تحتوي كلَ منها على عنصر واحد [5]، أو بمعنى أصح تهدف العنقدة الهرمية إلى الحصول على الشكل الهرمي للعناقيد المسماة بالمخطط الشجري (Dendrogram) الذي يبين عملية ارتباط العناقيد بعضها مع بعض من خلال سلسلة متداخلة من التجزيئات بإدماج العناقيد الصغيرة بصورة متكررة إلى عناقيد أكبر (طرائق التجميع) أو بفصل عناقيد كبيرة إلى عناقيد أصغر (طرائق التقسيم) كما هو مبين في الشكل (1) [1]، ويمكن لاساليب العنقدة الهرمية ان تجمع البيانات وتكوين العناقيد حسب نوعين الاول يبدأ بالعنصر بوصفة عنقوداً واحداً وفي كل خطوة يتم دمج أقرب زوج من العناصر في عنقود إلى أن يتم الوصول إلى عنقوداً واحداً يضم العناصر كافة ويسمى هذا الاسلوب طريقة العنقدة الهرمية المتكتلة (Agglomeration clustering) وهو الاكثر شيوعا وتحوي على العديد من الطرائق، اما الاسلوب الثاني فيختلف عن الاول في كونه يكون الشكل الهرمي بالترتيب معاكس اي تكون البيانات الكلية هي عبارة عن عنقود واحد ومن ثم في كل خطوة تتكون العناقيد المطلوبة وصولا الى عناقيد يحوي كل منها على عنصر واحد فقط [1].



شكل (1): يبين المخطط الشجري الهرمي التجميعي و التقسيمي للعناصر(A,B,C,D,E,F,G,H) ضمن مجموعة من العناقيد

2.1. طريقة وارد (Ward's Method) [1]:

أقترح هذه الطريقة العالم (Ward) في عام (1963) وهي تستخدم منهج تحليل التباين لتحديد المسافة بين العناقيد، أي تحاول التقليل من مجموع المربعات لأي عنقود من العنقودين المدمجين عند كل خطوة، وتعد هذه الطريقة هي الافضل من بين كل طرائق العنقدة المتكتلة والسبب في ذلك لان (Ward) افترض بأنه في كل مرحلة من مراحل التحليل العنقودي هناك كمية من المعلومات المفقودة والناتجة من عملية الدمج للعناقيد إذ يمكن حساب هذه المعلومات المفقودة من خلال حساب المجموع الكلي لمربعات انحرافات كل عنصر عن متوسط العنقود الذي ينتمي أليه وحسب الصيغة التالية:

$$E.S.S = \sum_{j=1}^{k}\left\{\sum_{i=1}^{n} xij^2 - \frac{\left(\sum_{i=1}^{nj} x_{ij}\right)^2}{n_j}\right\} \quad (1)$$

اذ ان $E.S.S$: تمثل المجموع الكلي لانحرافات القيم المربعة عن مركز العنقود الذي تنتمي أليه، و (X_i) هي العناصر او المفردات الموجودة في العنقود.

2.2. طريقة العنقدة المركزية (Centroid Clustering) [1]:

هي طريقة تعتمد على حساب المسافات بين المفردات لتكوين العناقيد، وتحسب المسافة بين عنقودين بالاعتماد على مراكز العناقيد:

$$D\ mean(s_i, s_j) = |mean_i - mean_j| \quad (2)$$

اذ (S_i, S_j) هما العنقودين المفترضين المراد حساب المسافة (D) بينهما، و ($mean_i$, $mean_j$) هما الوسطيين الحسابيين لهذيين للعنقودين والذان يحسبان وفق الصيغة الاتية:

ISSN (1681- 6870)

$$\text{mean}_i = \left(\frac{\sum_{i=1}^{n} x_i}{n}\right) \quad \& \quad \text{mean}_j = \left(\frac{\sum_{j=1}^{n} x_j}{n}\right) \tag{3}$$

حيث (X_i, X_j) هي العناصر الموجودة بكل عنقود، أما في حال استخدمنا الوسيط فالصيغة ستكون:

$$\left(D\ med(s_i, s_j) = |med_i - med_j|\right) \tag{4}$$

اذ ان (med_i, med_j) هما وسيطا العنقودين (S_i, S_j).

3. تحليل الانحدار (Linear Regression):

تحليل الانحدار يمكن أن يعرف بشكل عام كتحليل للعلاقة بين المتغيرات وهو احد الاساليب الإحصائية الأكثر استعمالا لأنه يعطينا طريقة سهلة لتحديد العلاقة بين المتغيرات، هذه العلاقة يمكن التعبير عنها بشكل معادلة تحتوي على متغير الاستجابة أو المتغير التابع (Y) مع واحد أو أكثر من المتغيرات المستقلة $(x_i, i = 1,2, \dots, n)$ [4]، أو أن تحليل الانحدار هو مجموعة الطرائق الإحصائية التي تتعامل مع الصيغ المختلفة للنماذج الرياضية التي تصف العلاقة بين المتغيرات بحيث يمكن استخدام نماذج هذه العلاقات لغرض التنبؤ والاستنتاجات الإحصائية الأخرى، وتحليل الانحدار يقوم على أنموذجين أساسيين هما (خطي، غير خطي)، ويعتبر الأنموذج الخطي من النماذج الأكثر شيوعا وهذا الأنموذج ينطوي على نوعين هما (بسيط، متعدد)، البسيط يقوم على العلاقة بين متغير معتمد أو تابع (Y) ومتغير مستقل واحد (X) وأما المتعدد فيقوم على العلاقة بين متغير تابع (Y) مع أكثر من متغير مستقل (X_i) [4]، وفي كلتا الحالتين أن كان بسيطاً أو متعدداً فلغرض أيجاد التنبؤات لابد من توفر معادلة تقديرية يعتمد عليها أو تكون الأساس في الاختبارات والتعويضات ومثل هذا المعادلة يتم إيجادها بعد تقدير معالم الأنموذج الخاص بالانحدار وهو الأنموذج الذي يأخذ الشكل الاتي في حالة كان انموذج خطي بسيط:

$$Y_i = \beta_0 + \beta_1 X_i + U_i \qquad i = 1,2, \dots \dots, n \tag{5}$$

حيث (Y_i) تمثل المتغير التابع، و (X_i) هي المتغير المستقل، اما (β_0) و(β_1) هما على التوالي الحد الثابت والميل الحدي لمعادلة الانحدار، و(U_i) هو حد الخطأ لهذه المعادلة، وياخذ الانموذج الشكل الاتي في حالة كان خطي متعدد:

$$Y = X\beta + U \tag{6}$$

ويكون هذا التقدير بطريقتين أما بطريقة المربعات الصغرى (Least Squares Method) أو طريقة الإمكان الأعظم (Maximum Likelihood Method) [4]، وأن طريقة المربعات الصغرى هي الاكثر شيوعا وأساس هذه الطريقة يعتمد على حساب قيم المعالم المجهولة $(\beta_0 \text{ و } \beta_1)$ لنموذج الانحدار والتي تجعل $(\sum_{i=1}^{n} U_i)$ في نهايتها الصغرى حيث أن قيم المعالم المحسوبة بهذه الطريقة تسمى تقديرات المربعات الصغرى وعندها نحصل على معادلة الانحدار التقديرية الاتية:

$$\hat{Y}_i = b_0 + b_1 X_i \tag{7}$$

$$\hat{Y} = Xb + U \tag{8}$$

وتسمى معادلة الانحدار التقديرية أو المعادلة التنبؤية وتعتبر المعادلة الأساسية التي يتم على أساسها إيجاد التنبؤات المرجوة للسنوات المقترحة وذلك بعد تعويض قيم المفردات المدروسة لكل من المتغير التابع والمتغير المعتمد لإيجاد المعادلة التقديرية، اذ يتم تعويض تسلسل السنوات بدل كل متغير (X_i) في ذات المعادلة التقديرية فنحصل على قيم المتغير $(\hat{Y}_i)$ والذي يمثل القيم التنبؤية المطلوبة للمحاصيل المدروسة، و$(b_0 \text{ و } b_1)$ هما الحد الثابت والميل الحدي لمعادلة الانحدار التقديرية على التوالي والتي تحسب كالاتي:

$$b_0 = \bar{Y} - b_1 \bar{X} \tag{9}$$

$$b_1 = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2} \tag{10}$$

ISSN (1681- 6870)

4. الجانب التطبيقي

في هذه الفقرة استخدمت بيانات محاصيل الخضروات الأساسية (الطماطم، الخيار، الباذنجان، الباميا) المتمثلة بالمتغيرات (أجمالي المساحة المزروعة، متوسط غلة الدونم لأجمالي المساحة، الإنتاج) ولعامي (2005-2010) لغرض الوصول الى نتائج التصنيف المرجوة، أن جوهر أسلوب التحليل العنقودي وأساسه المعتمد في تفسير النتائج هي طبيعة وشكل العناقيد المتكونة بين المتغيرات المدروسة وهذه العناقيد تتكون بين نتيجة لتجانس هذه المتغيرات فيما بينها واختلافها مع متغيرات العناقيد الأخرى وكما وضح في الفقرات السابقة، في هذا البحث فان معيار تجانس المتغيرات في العناقيد كانت الأراضي الزراعية للمحاصيل المدروسة فبعد أن تم إدخال المتغيرات المذكورة أعلاه لهذه المحاصيل ظهرت العناقيد وكلاً حسب أسلوب التحليل المتبع على شكل مناطق زراعية متباينة ومقسمة أو بمعنى أصح صنفت على شكل أراضي زراعية حسب المساحة وكمية الإنتاج وهو الهدف المتوخى من هذا البحث وكل منطقة تتكون من مجموعة من المحافظات والتي غالباً ما تكون محافظات متجاورة وهذا ما يؤكد الطبيعة الطبوغرافية المقسمة للعراق، ونتائج تصنيف الاراضي الزراعية حسب كل طريقة معتمدة عرض كالاتي:

4.1.طريقة وارد للربط (Ward's method)

أ. لعام (2005):

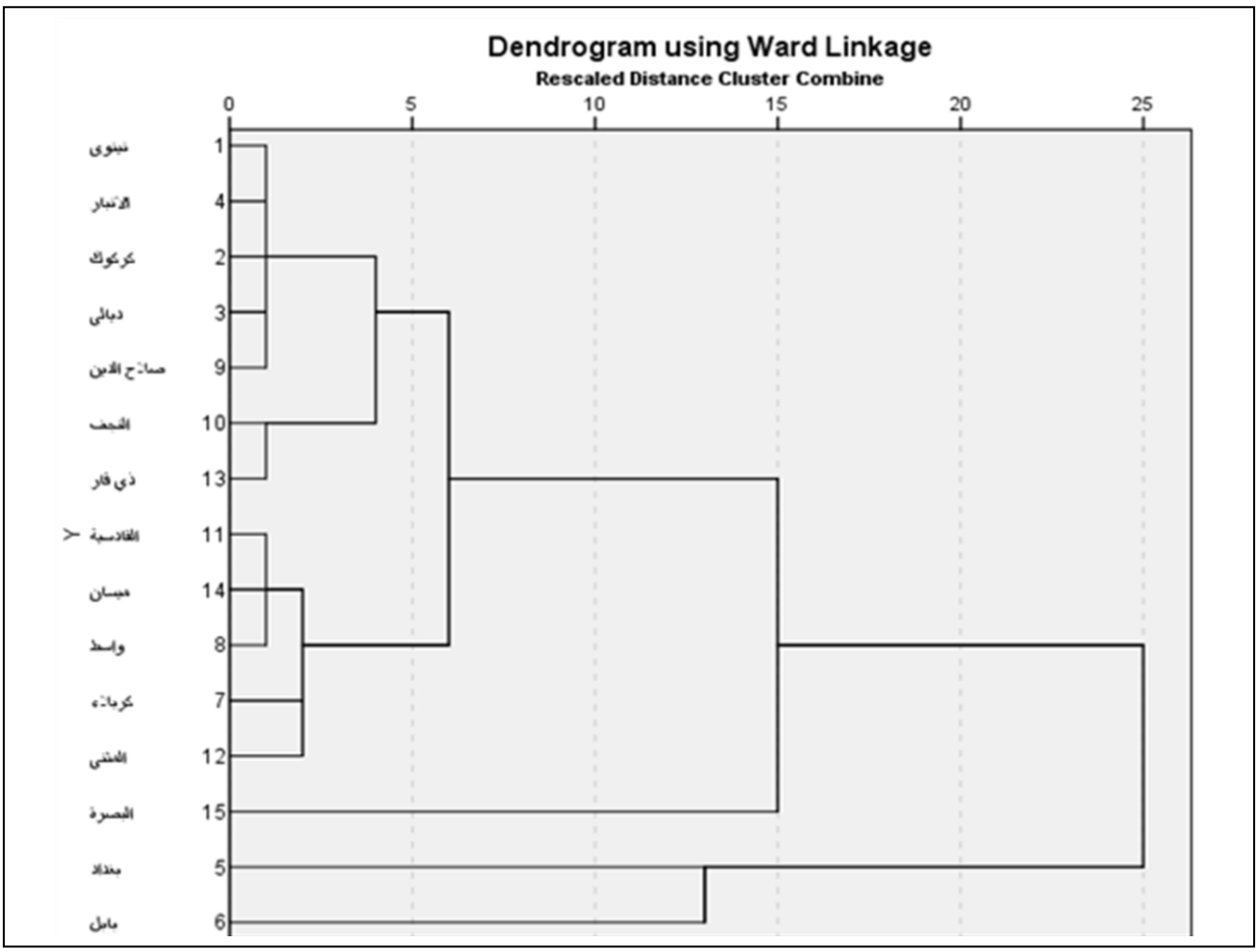
ومن خلال الاطلاع على جدول خطوات التجميع (Agglomeration Schedule Table) لعام (2005) والذي يعد الجزء الأهم في التفسير والذي منه يمكن تحديد المفردات أو المجموعات التي يتم ربطها في كل خطوة من خطوات التحليل يمكن ملاحظة ان عملية الربط بين المتغيرات تتم على أساس المسافات (المسافة الأقليدية) الفاصلة بينها أي أن المسافة الأقصر ستمثل المرحلة الأولى من العنقدة وأيضاً العناقيد هي عبارة عن تجمع مجموعة من المحافظات على أساس المساحات الزراعية المشتركة بينها، فنتائج الجدول بينت ان العنقود الأول تكون لأراضي الزراعية بين محافظة نينوى (المفردة 1) مع محافظة الأنبار (المفردة 4) بمسافة أقليدية قدرها (0.056) وهي الأقصر من غيرها وسيضاف أليهما متغير جديد في الخطوة الثالثة وهي محافظة كركوك (المفردة 2) بمسافة قدرها (0.213)، أما الخطوة الثانية فقد تم ربط محافظة القادسية (المفردة 11) بمحافظة ميسان (المفردة 14) بمسافة قدرها (0.130) وهكذا وصولا الى الخطوة الأخيرة التي تجمعت فيها محافظة نينوى (المفردة 1) مع بغداد (المفردة 5) بمقدار مسافة قدرها (11.002)، بالتالي فان مراحل التعنقد التي كما ذكرنا سابقاً ما هي ألا عناقيد بنيت على أساس الأراضي الزراعية المشتركة بصفات معينة بين متغيرات كل عنقود متكون، ونتائج الجدول جاءت كالاتي:

جدول (1): يبين خطوات التجميع حسب طريقة وارد للربط لعام (2005)

Agglomeration Schedule Table						
Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	1	4	.056	0	0	3
2	11	14	.130	0	0	4
3	1	2	.213	1	0	7
4	8	11	.330	0	2	9
5	3	9	.453	0	0	7
6	10	13	.596	0	0	10
7	1	3	.757	3	5	10
8	7	12	1.010	0	0	9
9	7	8	1.371	8	4	11
10	1	10	1.905	7	6	11
11	1	7	2.760	10	9	13
12	5	6	4.711	0	0	14
13	1	15	7.013	11	0	14
14	1	5	11.002	13	12	0

ولتفسير جدول خطوات التجميع السابق بشكل اكثر وضوحاً، يمكن معاينة المخطط الشجري (Dendogram) المستخرج للعام (2005) لملاحظة المفردات أو المجموعات التي تم ربطها معاً في كل خطوة من خطوات التحليل، فالتجميع أو التعنقد بين المشاهدات يعتمد على المسافة الأقصر بينهما وذلك لكونها أكثر تجانساً من الأزواج الأخرى من المشاهدات أي على سبيل المثال وكما يلاحظ من جدول خطوات التجميع السابق أو المخطط الشجر الاتي أن المحافظات الجنوبية (القادسية، ميسان، واسط) قد تجمعت معاً بعنقود فيما بينها وهو نفس الأمر بالنسبة للمحافظات الوسطى وهذا التجمع أعتمد على المسافة بينهما أي أنها متجانسة فيما بينها أكثر من المحافظات الأخرى وهذا يدل على تجانس أو تقارب طبيعة الزراعة المستخدمة فيها أو طبيعة التربة أو طبيعة الوسائل الزراعية المستخدمة فيها والذي يتضح من خلال بيانات المحاصيل وتقارب متوسطات الغلة والإنتاجية فيها، ثم بعد ذلك تبدأ العناقيد المستقلة في التجمع معاً لتصل إلى العنقود النهائي والذي انضمت أليه محافظة بابل.

ISSN (1681- 6870)



شكل (2): يعرض المخطط الشجري حسب طريقة وارد للربط لعام (2005)

ب. لعام (2010):

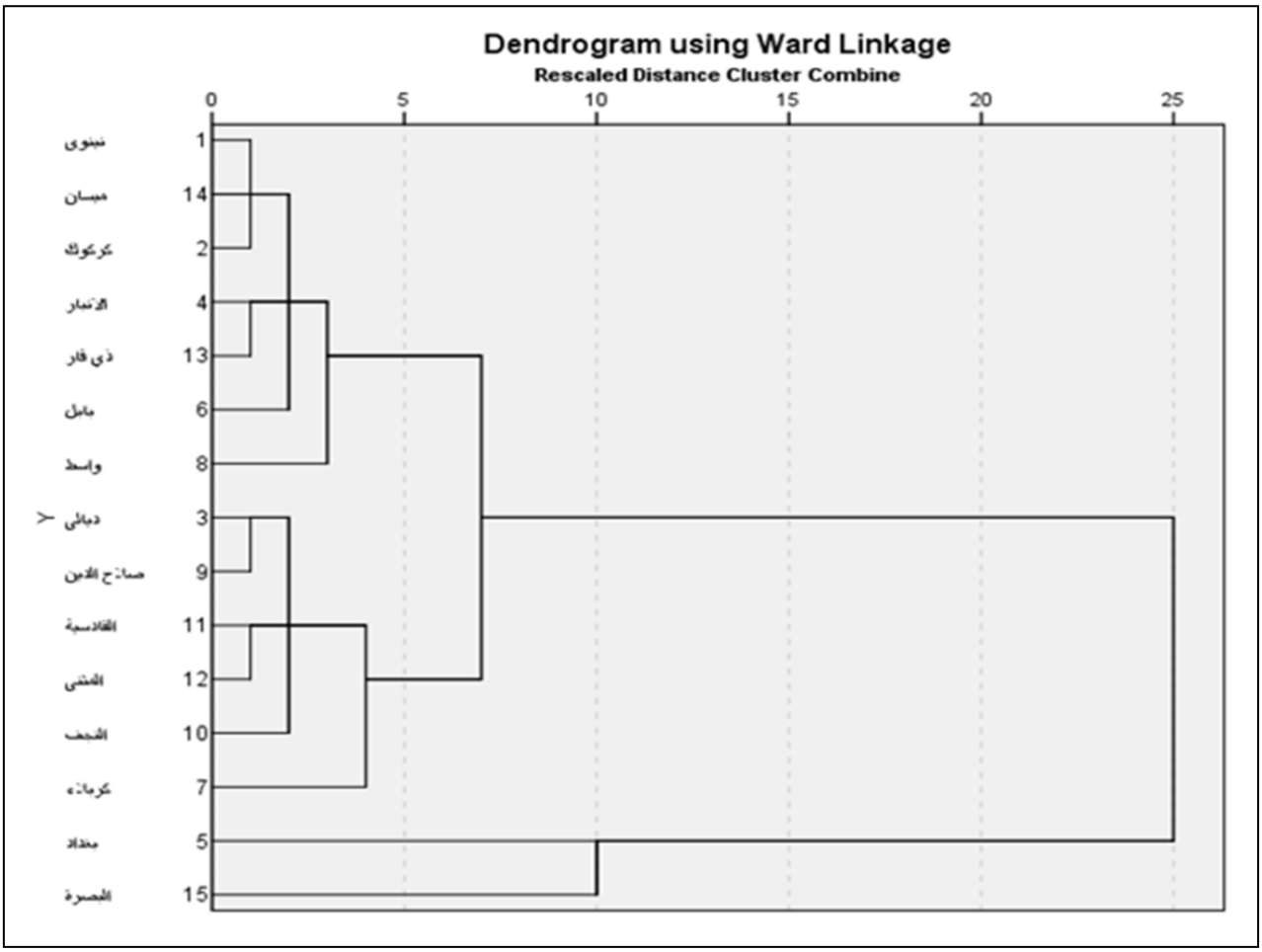
ونتائج جدول خطوات التجميع لهذا العام جاءت كالاتي:

جدول (2): يبين خطوات التجميع حسب طريقة وارد للربط لعام (2010)

Agglomeration Schedule Table						
Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	1	14	.045	0	0	2
2	1	2	.181	1	0	6
3	3	9	.319	0	0	7
4	11	12	.464	0	0	7
5	4	13	.643	0	0	6
6	1	4	.976	2	5	8
7	3	11	1.319	3	4	9
8	1	6	1.733	6	0	10
9	3	10	2.151	7	0	11
10	1	8	2.771	8	0	12
11	3	7	3.495	9	0	12
12	1	3	5.043	10	11	14
13	5	15	7.111	0	0	14
14	1	5	12.534	12	13	0

ISSN (1681- 6870)

يلاحظ أن في الخطوة الأولى تم تجميع محافظة نينوى (المفردة 1) مع ميسان (المفردة 14) بمسافة أقليدية قدرها (0.045) وهي الأقصر من غيرها وسيضاف أليهما متغير جديد في الخطوة الثانية وهو محافظة كركوك (المفردة 2) بمسافة قدرها (0.181) أما الخطوة الثالثة فجمعت بين محافظتي ديالى (المفردة 3) وصلاح الدين (المفردة 9) بمسافة قدرها (0.319) والخطوة الأخيرة ربطت المحافظتين نينوى (المفردة 1) مع بغداد (المفردة 5) بمقدار مسافة (12.534) وهكذا لبقية الخطوات أو مراحل التعنقد، وهو ما ظهر في المخطط الشجري لعام (2010) الاتي، اذ يلاحظ من الشكل تجمع محافظتي (نينوى، ميسان) بعنقود وأضيفت كركوك أليهم ثم أضيف أليهم عنقود (الانبار، ذي قار، بابل، واسط) وتستمر العناقيد بالتجمع وصولاً لإضافة عنقود (بغداد، بابل) وهذا ما يثبت صحة نتائج جدول خطوات التجميع السابق.



شكل (3): يعرض المخطط الشجري حسب طريقة وارد للربط لعام (2010)

4.2. طريقة العنقدة المركزية (Centroid clustering Method):

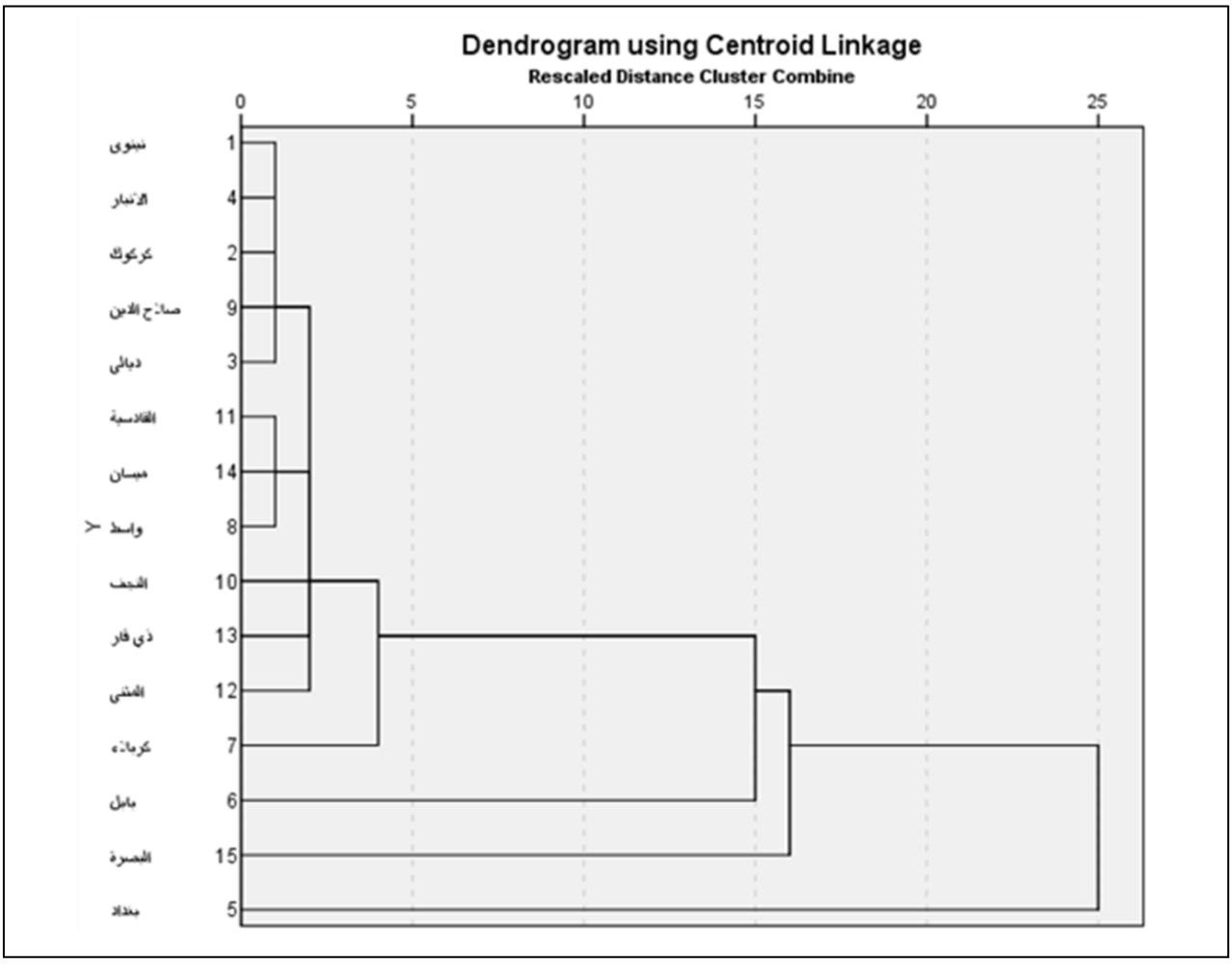
ونتائج التصنيف لهذا الاسلوب لكلا العامين (2005) و(2010) توضح أن محافظة نينوى (المفردة 1) كانت تتصدر جدول خطوات تجميع العناقيد اذ يبتدأ التعنقد من خلالها مع بقية المحافظات، ويمكن أن نلاحظ أيضاً أن هذه المحافظة كانت تتعنقد اولاً مع محافظات شمالية (كركوك، صلاح الدين) ومن ثم المحافظات الوسطى، أما في عام (2010) فكان عنقودها الأول يبدأ من محافظة جنوبية هي ميسان (المفردة 14)، وهذه النتائج موضحة أيضاً في المخططات الشجرية الخاصة بالعامين والتي توضح تعنقد محافظة نينوى مع نفس المحافظات المعروضة في جدول خطوات التجميع، وأخيراً يلاحظ بأن الاعضاء المكونين للعناقيد (المحافظات المتجمعة في كل عنقود) والمعروضة في جداول خطوات التجميع هم نفس الاعضاء لكلا العامين لكن باختلاف تسلسل البدء بالتعنقد لاختلاف البيانات بين السنتين المعتمدتين وبالتالي المسافات الأقليدية بين العناصر المكونة للعناقيد، وعرضت نتائج جداول خطوات تجميع العناقيد والمخططات الشجرية لكلا العامين (2005) و(2010) حسب الترتيب الاتي:

ISSN (1681- 6870)

أ. عام (2005):

جدول (3): يبين خطوات التجميع حسب طريقة العنقدة المركزية لعام (2005)

Agglomeration Schedule Table						
Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	1	4	.111	0	0	2
2	1	2	.125	1	0	4
3	11	14	.149	0	0	5
4	1	9	.166	2	0	6
5	8	11	.175	0	3	7
6	1	3	.199	4	0	7
7	1	8	.276	6	5	9
8	10	13	.287	0	0	9
9	1	10	.364	7	8	10
10	1	12	.367	9	0	11
11	1	7	.621	10	0	12
12	1	6	2.329	11	0	13
13	1	15	2.531	12	0	14
14	1	5	4.010	13	0	0



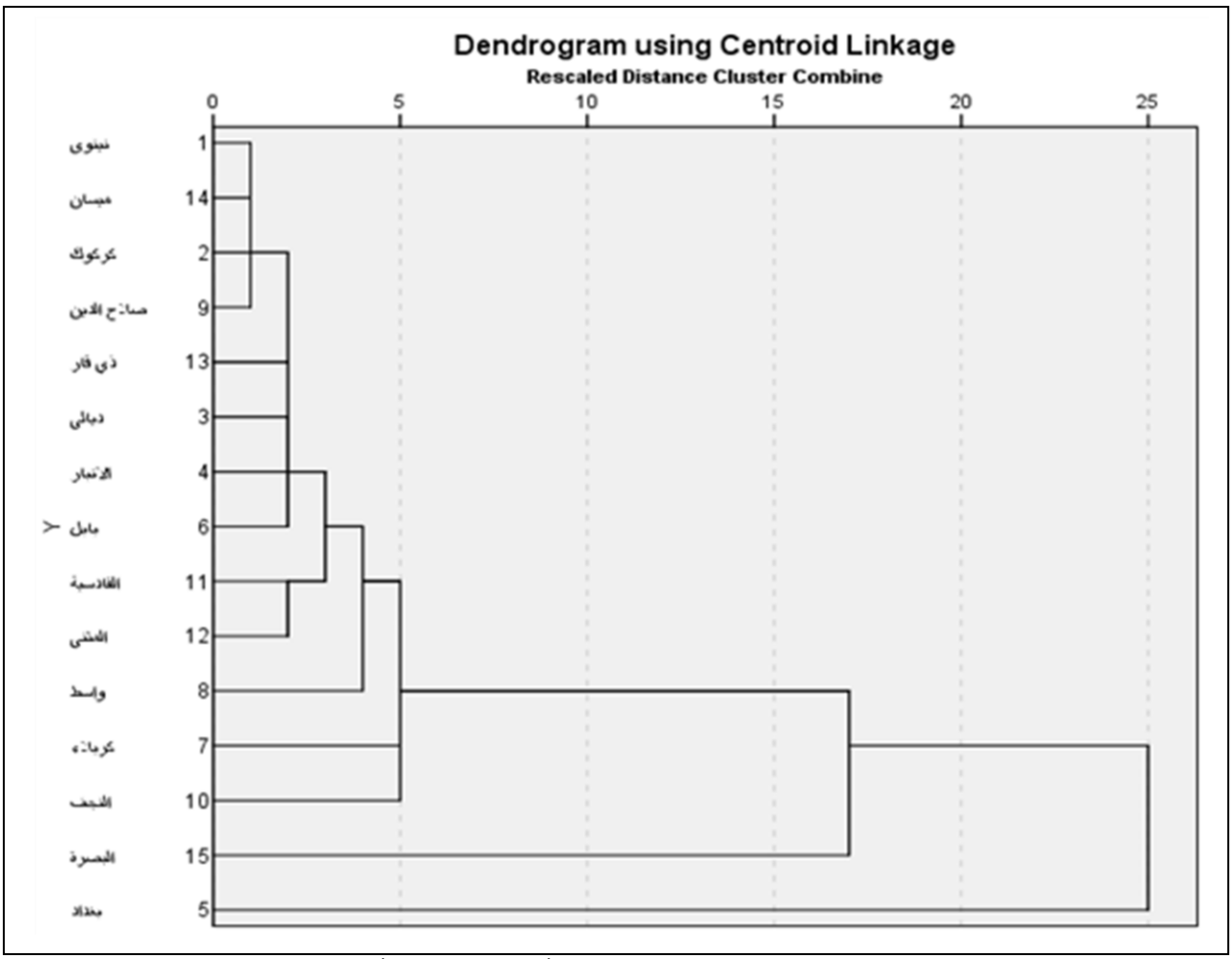
شكل (4): يعرض المخطط الشجري حسب طريقة العنقدة المركزية لعام (2005)

ISSN (1681- 6870)

ب. عام (2010):

جدول (4): يبين خطوات التجميع حسب طريقة العنقدة المركزية لعام (2010)

Agglomeration Schedule Table						
Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	1	14	.091	0	0	2
2	1	2	.204	1	0	3
3	1	9	.267	2	0	5
4	11	12	.290	0	0	9
5	1	13	.309	3	0	6
6	1	3	.366	5	0	7
7	1	4	.391	6	0	8
8	1	6	.455	7	0	9
9	1	11	.497	8	4	10
10	1	8	.756	9	0	11
11	1	7	.869	10	0	12
12	1	10	1.032	11	0	13
13	1	15	3.231	12	0	14
14	1	5	4.811	13	0	0



شكل (5): يبين المخطط الشجري حسب طريقة العنقدة المركزية لعام (2010)

ISSN (1681- 6870)

وبالتالي ومن خلال نتائج طرائق التحليل الهرمي نلاحظ بان هذه الاساليب كانت متقاربة في نتائجها فكلها أجمعت على بدء محافظة نينوى بالتعنقد ولكلا العامين وتعنقد هذه المحافظة كان غالباً ما يكون مع محافظات شمالية (كركوك أو صلاح الدين)، وكما هو معلوم فان مقياس المسافة هو المؤشر على مدى تجانس المفردات وما بدء التعنقد من المحافظات الشمالية إلا دليل على تجانسها اذ كلما كانت مقاييس المسافة قصيرة بين العناقيد كان هذا يدل على تجانس مفردات العنقود، ومن ثم تبدأ بالتعنقد مع محافظات الوسط والتي بدورها كانت قد تعنقدت مع محافظات الجنوب ومن ثم تجمعت العناقيد مكونة العنقود النهائي لكن رغم تعنقد محافظات الوسط مع الجنوب إلا إن البدء بالعنقدة كان من الشمال لأنه صاحب المسافة الاقليدية الأقصر وهو الأمر الواضح في جداول التجمع وبشكل أوضح بالمخططات الشجرية، وأيضا يلاحظ تقارب نتائج جداول أعضاء العناقيد الا في بعض الفروقات البسيطة.

4.3. التنبؤ (Production)

بعد أن قمنا بتفسير النتائج المستحصل عليها نأتي الآن إلى جداول التنبؤات المستخرجة لإنتاجية المحاصيل في العراق، كما ذكر سابقا فمن أجل الحصول على التنبؤات لابد من توفر معادلة تقديرية يتم التنبؤ على أساسها وكيفية الحصول أو الوصول لمثل هذه المعادلة تطرقنا أليها في الفقرة (3) السابقة وبالتالي استطعنا الحصول على المعادلات التقديرية الخاصة بكل محصول زراعي ولانتاجية العراق وذلك بعد تعويض القيمة الإجمالية لمتغيرات الإنتاج كمتغير معتمد (Y_i) واجمالي المساحة كمتغير مستقل (X_i) حيث ظهرت المعادلات التقديرية كالتالي:

المعادلة التقديرية لمحصول الطماطم في العراق:

$$\hat{Y}i = 578130.684 + 1.556\ Xi \qquad (11)$$

المعادلة التقديرية لمحصول الخيار في العراق:

$$\hat{Y}i = -5591.632 + 2.296\ Xi \qquad (12)$$

المعادلة التقديرية لمحصول الباذنجان في العراق:

$$\hat{Y}i = 15552.408 + 2.764\ Xi \qquad (13)$$

المعادلة التقديرية لمحصول البامية في العراق:

$$\hat{Y}i = 48578.281 + 1.167\ Xi \qquad (14)$$

حيث ظهرت نتائج التنبؤ التقريبية كما في الجدول التالي:

جدول (5): يبين القيم التنبؤية لكميات انتاج محاصيل الخضراوات في العراق

السنة	التسلسل	الطماطم	الخيار	الباذنجان	الباميا
2011	7	578142	5576 -	15572	48586
2012	8	578143	5573 -	15575	48588
2013	9	578145	5571 -	15577	48589
2014	10	578146	5569 -	15580	48590
2015	11	578148	5566 -	15583	48591
2016	12	578149	5564 -	15586	48592
2017	13	578151	5562 -	15588	48593
2018	14	578152	5559 -	15591	48595
2019	15	578154	5557 -	15594	48596
2020	16	578156	5553 -	15596	48597

يلاحظ أن السنوات أعلاه ابتدأ تسلسلها من الرقم (7) وذلك لكون البيانات المستخدمة في التسلسل ابتدأت من السنة (2005) وكان تسلسلها (1) والسنة الأخيرة المعلومة هي (2010) وتسلسلها (6)، علماً أن الإنتاج المتنبئ به بالأطنان، والنتائج أعلاه تبين أن محاصيل (الطماطم، الباذنجا، الباميا) كانت كميات أنتاجها متزايدة على الرغم من أن هذه الكميات هي تقريبا نصف الكميات المنتجة في العام الأخير (2010)، أما محصول الخيار فقد كان متناقصاً وربما يعود هذا التفاوت في كميات الإنتاج إلى قصر السلسلة الزمنية المستخدمة في التنبؤ، ويمكن أيضاً استخراج القيم التنبؤية للمحافظات الأعلى إنتاجاً لهذه المحاصيل والتي كانت معادلاتها التقديرية كالاتي:

ISSN (1681- 6870)

المعادلة التقديرية لمحصول الطماطم في محافظة البصرة:

$$\hat{Y}i = 50906.444 + 6.3761\ Xi \quad (15)$$

المعادلة التقديرية لمحصول الخيار في محافظة بغداد:

$$\hat{Y}i = -235.602 + 2.129\ Xi \quad (16)$$

المعادلة التقديرية لمحصول الباذنجان في محافظة بغداد:

$$\hat{Y}i = 43485.165 + 6.149\ Xi \quad (17)$$

المعادلة التقديرية لمحصول البامية محافظة بغداد:

$$\hat{Y}i = -2878.406 + 2.871\ Xi \quad (18)$$

حيث ظهرت النتائج التنبؤ التقريبية كما في الجدول التالي:

جدول (6): يبين القيم التنبؤية لكميات انتاج محاصيل الخضراوات في المحافظات

السنة	التسلسل	الطماطم	الخيار	الباذنجان	البامية
2011	7	50951	-221	43528	-2858
2012	8	50957	-219	43534	-2855
2013	9	50964	-216	43541	-2852
2014	10	50970	-214	43547	-2850
2015	11	50977	-212	43553	-2847
2016	12	50983	-210	43559	-2844
2017	13	50989	-208	43565	-2841
2018	14	50996	-206	43571	-2838
2019	15	51002	-204	43577	-2835
2020	16	51008	-202	43583	-2832

ويلاحظ من النتائج أعلاه محصول الطماطم هو الوحيد المتزايد والمحاصيل الأخرى متناقصة في أنتاجها وربما يعود هذا التفاوت إلى قصر السلسلة التنبؤية أو قلة الأراضي المستغلة في الزراعة في هذه المحافظات، أن النتائج التي ظهرت بالإشارة السالبة تعني أن الإنتاج متناقص أو هو متناقص وقد بدأ بالتزايد إذ لايوجد إنتاج سالب وأيضا يلاحظ أن النتائج كانت متقاربة ويعود السبب إلى قصر السلسلة المستخدمة وأيضاً أن تحليل الانحدار صحيح انه احد الأساليب المستخدمة للتنبؤ لكن عادة ما يعتمد هذا الأسلوب في تقدير قيمة أو التنبؤ لقيمة واحدة مفقودة من سلسلة.

5. الاستنتاجات

1. إن أفضل المناطق صلاحية للزراعة هي المناطق الشمالية (كركوك، نينوى، صلاح الدين) وتليها محافظات الوسط والجنوب وذلك لكون تسلسل العناقيد وبدئها بالتعنقد كان دائماً ما يبدأ من هذه المحافظات، وإن أفضل المحافظات صلاحية للزراعة كانت محافظة نينوى وذلك لكونها تأتي دائماً كأساس في تجميع العناقيد مع بعضها البعض.
2. طرائق التحليل العنقودي الهرمي المختلفة كانت متقاربة في نتائجها وهذا الأمر كان واضحاً في جداول خطوات التجميع حيث كانت المفردات المكونة للعناقيد هي نفسها لكن باختلاف تسلسل بدأ تعنقد هذه العناقيد وهذا الاختلاف يعود إلى اختلاف المعادلة الخاصة بحساب المسافة الأقليدية في كل طريقة او للتفاوت في الكميات المنتجة من محاصيل الخضر.
3. مجمل المحاصيل الزراعية وحسب نتائج التنبؤ اتضح ان إنتاجها متزايد مع الأعوام عدا محصول الخيار حيث ظهرت معادلته التقديرية سالبة مما أدى إلى أن يكون إنتاج الخيار متناقصاً لكن مع ذلك يمكن ملاحظة أن إنتاج المحاصيل في عام (2011) متزايداً كان أم متناقصاً هو اقل بمقدار النصف عما هو علية في عام (2010) وربما يعود سبب هذا التناقص إلى قصر السلسلة المستخدمة في التنبؤ فكما هو معلوم أن التنبؤ يعتمد على تعويضات تسلسل السنوات المراد التنبؤ لها .

6. التوصيات

1. محاولة استخدام أساليب أخرى معتمدة في التصنيف كأسلوب التحليل المميز وذلك لغرض المقارنة بين أساليب التصنيف وطبيعة استنتاجاتها لغرض الاستفادة من هذه الاستنتاجات في تطوير الزراعة في العراق وايضا السعي لدراسة محاصيل زراعية أكثر من أجناس أخرى غير الخضروات وذلك لمعرفة مقدرة المحافظات العراقية على انتاج محاصيل مختلفة وليس محصول معين والذي يسهم بدوره بالاكتفاء الذاتي.

ISSN (1681- 6870)

2. اعتماد سلاسل بيانية أطول في إيجاد التنبؤات المستقبلية للمحاصيل وذلك لكي تكون النتائج أكثر دقة وملائمة لكي تعتمد في الخطط المستقبلية المتعلقة بدراسة وتطوير إنتاجية المحاصيل الزراعية، وايضا محاولة استخدام أساليب إحصائية مختلفة لإيجاد التنبؤات المستقبلية كأسلوب السلاسل الزمنية مثلاً والمقارنة فيما بينها من اجل اعتماد الأسلوب الأفضل في الخطط المستقبلية وذلك لكون تحليل الانحدار عادة ما يستخدم لإيجاد تنبؤات قيم محددة اكثر من استخدامه في إيجاد تنبؤات سلسلة من السنوات.
3. زيادة الأراضي المستغلة زراعياً إما عن طريق استصلاح أراضي لم تستصلح سابقاً أو التقليل من التمدد العمراني أو مكافحة التصحر وتقليص امتداده باتجاه المناطق الزراعية ومحاولة طرح الأراضي الغير مستصلحة البعيدة عن التجمعات السكانية أو الصعبة الاستصلاح على شركات استثمارية متخصصة في مثل هكذا ظروف لغرض الاستفادة من عائدات هذه الأراضي واستثمارها في دعم الزراعة في مناطق أخرى أو في دعم ميزانية الدولة.
4. دعم المنتج المحلي من خلال توفير أسباب نجاح زراعته وأيضاً محاولة تسهيل تسويقه في الأسواق المحلية وذلك من خلال دعم وتوحيد أسعار المنتجات الزراعية لغرض منافسة المنتجات المستوردة وايضا دعم الفلاح بما يحتاجه من وسائل تساعده في الزراعة من أسمدة وبذور محسنة ووسائل ري حديثة ومكائن حرث وبذار ومحاولة تخصيص نسب ثابتة من مياه الري لكل محافظة لغرض التقليل من حالات التجاوز على كميات مياه الري المخصصة لكل محافظة، ومحاولة تشريع قوانين تساهم في تنضم عملية الاستيراد للمنتجات الزراعية وذلك من خلال وضع بعض القيود الـگمرگية أو الضرائب أسوة بباقي الدول الزراعية المنتجة التي تسعى لحماية منتجها الوطني.

المصادر

[1] بيثون، نغم نافع "التحليل العنقودي لاستخراج القيم الشاذة مع تطبيق عملي"، أطروحة دكتوراه في الإحصاء، كلية الإدارة والاقتصاد، جامعة بغداد (2002).
[2] جبار، كاظم جبار "التحليل العنقودي وتطبيقاته العملية في تشخيص مرض الـگـلازار الذي يصيب الأطفال"، رسالة ماجستير في الإحصاء، كلية الإدارة والاقتصاد، جامعة بغداد (1989).
[3] جوني، أمجد سامي "استخدام التحليل العنقودي في تصنيف الأسر حسب مستويات المعيشة للفترة (2007-2006)"، بحث دبلوم إحصاء، كلية الإدارة والاقتصاد، جامعة بغداد (2011).
[4] عبد الحميد، أسيل عوض "تقدير وتصنيف المرونات الأنفاقية لبعض المجاميع السلعية في العراق باستخدام نماذج الانحدار والتحليل العنقودي"، بحث دبلوم إحصاء، كلية الإدارة والاقتصاد، جامعة بغداد (2012).
[5] ناصر، عبير عباس "استعمال التحليل العنقودي لبيان تغير نمط الأنفاق للفترة من (1971-2010)"، أطروحة دكتوراه في الإحصاء، كلية الإدارة والاقتصاد، جامعة بغداد (2012).