

مقارنة جودة بعض طرق التنبؤ قصيرة الأجل ببيانات السلسلة الزمنية الهرمية لإنتاج القمح الطري في سورية

فلك السطايحي * (1) ومحمد ظاهر عنان (2) وأمجد ماسو (2)

(1). قسم العلوم الأساسية، كلية الهندسة الكيميائية والبتروولية، جامعة البعث، حمص، سورية.

(2). قسم الإحصاء الرياضي والبرمجة، كلية العلوم، جامعة حلب، حلب، سورية.

(*المراسلة الباحثة: فلك السطايحي، البريد الإلكتروني: falaksataihi@yahoo.com)

تاريخ القبول: 2021/07/4

تاريخ الاستلام: 2021/03/31

الملخص

قام البحث بدراسة طرق توليد التنبؤات بالسلاسل الزمنية الهرمية على المدى القصير وتطبيقها على بيانات إنتاج القمح الطري في سورية والتي صنفت حسب المنطقة ومن ثم حسب المدينة لسلاسل زمنية من 2006–2018. حيث نوقشت أربع طرق وهي Top-Down، Bottom-Up، Middle-Out و Optimal Combination، وطبقت إحدى معايير اختبار جودة طرق التنبؤ وهي MASE لمعرفة الطريقة الأفضل أداء لاعتمادها في الأبحاث ذات الصلة بهذا البحث من أجل تسهيل وتسريع إنجاز العمليات الحسابية. وقد بينت النتائج أن جميع الطرق الأربعة تختلف عن بعضها بشكل عام. ونتيجة للدراسة التجريبية تبين أن أداء طريقة Top-Down أفضل من أداء بقية الطرق في التنبؤ، وتم التنبؤ لثلاث سنوات مستقبلية بعد 2018 حيث أشارت النتائج بتوقع الزيادة في إنتاج محصول القمح الطري على جميع مستويات التركيب الهرمي للسلسلة المدروسة.

الكلمات المفتاحية: السلاسل الزمنية الهرمية، Top-Down، Bottom-Up، Middle-Out، Optimal Combination.

المقدمة:

يعد القمح من أقدم المحاصيل الزراعية التي عرفها الإنسان ولا يزال هذا المحصول من أكثرها انتشاراً، يزرع القمح بنوعيه القاسي (Durum Wheat) وهو الأكثر شيوعاً والقمح الطري (Soft Wheat) في مساحات واسعة مروياً أو بعللاً، وتتباين كمية الانتاج تبعاً لعوامل عديدة، من أهمها الظروف المناخية خلال الموسم الزراعي وبخاصة كميات الهطول المطري وتوزعه، حيث بلغ إنتاج القمح في سوريا حوالي 1.2 مليون طن بمتوسط إنتاجية نحو 1115 كغ للهكتار الواحد من النوعين الطري والقاسي مع مساحة الأراضي المزروعة بالقمح قرابة 1.1 مليون هكتار، مثلت نسبة إنتاج القمح الطري حوالي 38%، ونسبة المساحة حوالي 44% بمتوسط إنتاجية للهكتار الواحد نحو 969/كغ، وتأتي كل من محافظة الحسكة وحلب والرقبة بالمرتبة الأولى والثانية والثالثة بالإنتاج بنسب مثلت نحو 34%، 26%، 21% من إجمالي إنتاج القمح الطري على الترتيب (وزارة الزراعة والإصلاح الزراعي، 2018).

يتم في العديد من البحوث الزراعية والاقتصادية إنشاء نموذج تنبؤ يمثل الظاهرة المتاحة ويمكن من الحصول على تنبؤات تساعد في إعداد النمذجة والتخطيط. باعتبار أن عملية التنبؤ هي عملية تقدير لما سيحدث مستقبلاً لظاهرة ما اعتماداً على تحليل بياناتها في الماضي باستخدام نموذج رياضي مناسب. وتعمل التنبؤات الدقيقة على رسم السياسات المستقبلية في كافة المجالات، إذ تعمل على تقدير مسار الظاهرة في المستقبل ومعدل نموها والتغيرات المصاحبة لها قبل اتخاذ أي قرار بشأنها، لذلك يعتبر التنبؤ وسيلة لتحقيق عملية التخطيط للمستقبل على الوجه الأكمل من خلال الاستفادة من التنبؤات لتوجيه الظروف الحالية وفقاً للأهداف المراد تحقيقها في ضوء الإمكانيات المتاحة. وهناك نوعان من طرق التنبؤ: طرق تنبؤ نوعية وطرق تنبؤ كمية (جابر ومجيد 2017). وتعد دراسة السلاسل الزمنية طريقة من طرق التنبؤ الكمي تهدف إلى:

وصف الظاهرة محل الدراسة وتفسير سلوكها وفقاً لمشاهداتها عبر الزمن.

التنبؤ بالظاهرة محل الدراسة لفترة زمنية قصيرة المدى أو بعيدة المدى.

وفي ظل الحاجة المستمرة للأساليب الدقيقة التي تمكن من توليد تنبؤات تكون قريبة من الواقع ومن ثم تخدم صانعي القرار، فإن البحث يقوم على دراسة طرق التنبؤ بالسلاسل الهرمية ومعالجتها بحيث تؤدي أفضل دقة ممكنة.

ويوجد عدد من الدراسات ذات القيمة العالية والتي تناولت المناهج الحديثة بالتنبؤ للسلاسل الهرمية على المدى القصير ومنها:

في دراسة قام بها (Simon et al., 2006) قام البحث بدراسة إمكانات نماذج السلاسل الزمنية الهرمية لاكتشاف الانحرافات عن المعدلات المتوقعة التي قد تشير إلى تفشي الأمراض المعدية حتى تتخذ الإجراءات المناسبة للسيطرة على تفشي تلك الأمراض حيث تطلبت الكمية الهائلة من البيانات التي تم جمعها باستخدام أنظمة المراقبة الآلية القائمة على المختبرات خوارزميات آلية للكشف عن الانحرافات غير المتوقعة.

في دراسة قام بها (Hyndman and Khandakar, 2007) تناولت الدراسة وصف خوارزميتين للتنبؤ التلقائي حيث أنه غالباً ما تكون هناك حاجة إلى التنبؤات التلقائية لأعداد كبيرة من السلاسل الزمنية أحادية المتغير في الأعمال وسياقات أخرى. وقد تم تنفيذ الخوارزميتين في الحزمة forecast. الخوارزميات قابلة للتطبيق على كل من البيانات الموسمية وغير الموسمية، وتمت مقارنتها وتوضيحها باستخدام أربع سلاسل زمنية حقيقية.

في دراسة قام بها (Athanasopoulos et al., 2009) ناقش البحث خمس طرقاً للتنبؤ الهرمي: نوعان مختلفان من طرق Top-down وطريقة Bottom-up، ونهج من أعلى إلى أسفل مقترح حديثاً حيث يتم تصنيف تنبؤات المستوى الأعلى وفقاً للنسب المتوقعة لسلاسل المستوى الأدنى، ومؤخراً نهج التجميع الأمثل المقترح.

في دراسة قام بها (Hyndman et al., 2011) تم اقتراح طريقة إحصائية جديدة للتنبؤ بالسلسلة الزمنية الهرمية، والتي تتيح إنتاج تنبؤات مثالية أفضل من التنبؤات التي تنتجها طريقة Top-down أو Bottom-up. تعتمد تلك الطريقة على التنبؤ بشكل مستقل بكافة السلاسل على جميع مستويات التسلسل الهرمي، ثم استخدام نموذج الانحدار لدمج هذه التوقعات وتوفيقها على النحو الأمثل. تضاف التنبؤات الناتجة بشكل مناسب عبر التسلسل الهرمي، وهي غير منحازة ولديها أدنى تباين بين جميع تنبؤات المجموعات وفقاً لبعض الافتراضات البسيطة.

في الدراسة التي قام بها (Hyndman et al., 2016) توفر الخوارزميات الحسابية التي تمت مناقشتها طريقة للتعامل مع مجموعات هائلة من السلاسل الزمنية في حالتين خاصتين ولكن مهمتين (عندما يتم ترتيب السلاسل الزمنية في التسلسل الهرمي، وعندما يكون للسلسلة الزمنية هيكل تجميع ثنائي الاتجاه). وقد أظهر البحث كيفية الاستفادة من بنية قيود التجميع لتقليل التعقيد الحسابي للمشكلة.

في دراسة قام بها (Wickramasuriya et al., 2019) تم اقتراح نهجاً جديداً لتوفيق التنبؤات للسلاسل الزمنية الهرمية والمجموعة والذي يتضمن معلومات من كامل مصفوفة التغيرات لأخطاء التنبؤ وكذلك أظهرت الدراسة أن تطبيق التوفيق يضمن تنبؤات متسقة والتي تكون جيدة مثل التنبؤات الأساسية.

هدف البحث:

يهدف البحث إلى إيجاد أفضل طريقة للتنبؤ ببيانات السلاسل الزمنية الهرمية لإنتاج القمح الطري في سورية على المدى القصير لتطبيقها في الدراسات التي تخص هذا النوع من البيانات من أجل تسهيل وتسريع إنجاز العمليات الحسابية وذلك من خلال:

1. استخدام الطريقة التصاعدية Bottom-Up، الطريقة التنازلية Top-Down، الطريقة الوسطى Middle-Out.
2. استخدام الطريقة الحديثة Optimal Combination لتوليد التنبؤات بالسلاسل الزمنية الهرمية.

المفاضلة بين المنهجية في الفقرة 2 والطرق التقليدية للتنبؤ بالسلاسل الهرمية والمبتكرة من قبل العالم R. Hyndman مواد البحث وطرائقه:

- البيانات: اعتمد البحث على البيانات الثانوية المنشورة والصادرة عن وزارة الزراعة والإصلاح الزراعي في سورية لإنتاج القمح الطري على مستوى سورية والمناطق والمدن لسلاسل زمنية من 2006-2018.

- أسلوب التحليل: اعتمد البحث على بعض طرق التنبؤ قصيرة الأجل للسلاسل الهرمية (Optimal Combination، Middle-Out، Top-Down، Bottom-Up)، وتطبيق معيار (MASE) للمفاضلة بين الطرق المختلفة لجودة التنبؤ.

- البرنامج المستخدم في بناء السلسلة الهرمية وتنفيذ طرق التنبؤ:

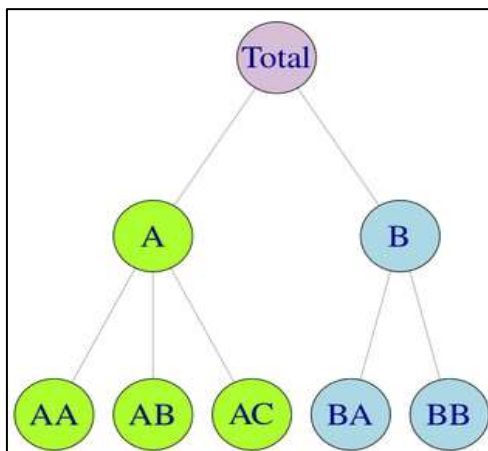
R Language (Package hts and forecast for Forecasting Hierarchical)

- المفاهيم والمصطلحات المتعلقة:

السلسلة الزمنية الهرمية Hierarchical time series:

تتكون السلسلة الزمنية الهرمية من سلاسل زمنية منظمة بشكل هرمي (Athanasopoulos et al., 2009) ويمكن تجميعها في مجموعات على عدة مستويات مختلفة بناءً على المنتجات أو الجغرافيا أو بعض الميزات الأخرى، تعطى تسمية لكل سلسلة من خلال سلسلة من الحروف تحدد وفقها السلسلة الفردية ومستوى التصنيف. كما يرمز إلى قيمة السلسلة الاجمالية Total في اللحظة t بواسطة y_t ويرمز إلى قيمة السلسلة في اللحظة t عند العقدة j بـ $y_{j,t}$. على سبيل المثال، يشير $y_{A,t}$ إلى قيمة السلسلة المقابلة للعقدة A في المستوى 1 عند اللحظة t ، كما أن $y_{AB,t}$ يشير إلى قيمة السلسلة المقابلة للعقدة AB في المستوى 2 عند اللحظة t ، وهكذا.

يوضح الشكل (1) تسلسل هرمي بمستويين حيث يكون الإجمالي " Total " في قمة التسلسل الهرمي (الذي نسميه المستوى صفر) وهو المستوى الأكثر تجميعاً للبيانات. يقسم الإجمالي إلى سلسلتين في المستوى 1، والتي تنقسم بدورها إلى ثلاث سلاسل وسلسلتين على التوالي في المستوى الأدنى من التسلسل الهرمي الذي يحتوي على السلاسل الزمنية الأكثر تصنيفاً.



الشكل (1): مخطط شجرة لتسلسل هرمي بمستويين

في هذا المثال الصغير، يكون العدد الإجمالي للسلاسل في التسلسل الهرمي هو $n = 1 + 2 + 5 = 8$ في حين أن عدد السلاسل في المستوى الأدنى هو $m = 5$. يلاحظ أن $n > m$ في كل تسلسل هرمي. كما أنه عند أي لحظة t ، تكون قيمة السلسلة في المستوى الأعلى من التسلسل الهرمي مساوية لمجموع قيم السلسلة في المستوى الأدنى. فمثلاً:

$$y_t = y_{AA,t} + y_{AB,t} + y_{AC,t} + y_{BA,t} + y_{BB,t} \quad (1)$$

$$y_{A,t} = y_{AA,t} + y_{AB,t} + y_{AC,t} \quad \text{و} \quad y_{B,t} = y_{BA,t} + y_{BB,t} \quad (2)$$

بتعويض (2) في (1) نحصل أيضاً على: $y_t = y_{A,t} + y_{B,t}$ يمكن اعتبار هذه المعادلات بمثابة قيود تجميع، ويمكن تمثيلها بشكل أكثر كفاءة باستخدام التمثيل المصفوفي حيث نقوم ببناء مصفوفة من المرتبة $n \times m$ تدعى بمصفوفة الجمع "Summing Matrix" والتي تحدد الطريقة التي يتم بها تجميع سلسلة المستوى الأدنى حيث أن السطر العلوي هو متجه واحد طوله m والجزء السفلي هو المصفوفة الواحدة من المرتبة $m \times m$.

الأجزاء الوسطى من S هي متجه مصفوفات مستطيلة قطرية. وبالتالي يمكن كتابة البنية هرمية في الشكل (1) كما يلي:

$$\begin{bmatrix} y_t \\ y_{A,t} \\ y_{B,t} \\ y_{AA,t} \\ y_{AB,t} \\ y_{AC,t} \\ y_{BA,t} \\ y_{BB,t} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_{AA,t} \\ y_{AB,t} \\ y_{AC,t} \\ y_{BA,t} \\ y_{BB,t} \end{bmatrix}$$

$$y_t = S b_t \quad \text{أو بشكل أكثر اختصاراً:} \quad (3)$$

حيث y_t عبارة عن متجه بعده n يحوي على كل قيم السلسلة الهرمية في اللحظة t و S هي مصفوفة الجمع، و b_t هو متجه بعده m يحوي كل القيم في المستوى الأدنى للتسلسل الهرمي في اللحظة t . من الواضح أن السطر الأول في مصفوفة الجمع S يمثل المعادلة (1) أعلاه، ويمثل الصفين الثاني والثالث المعادلة (2). تشتمل الأسطر التالية على مصفوفة واحدة m بعدها m ، بحيث كل قيمة من المستوى السفلي في الجانب الأيمن من المعادلة تساوي نفس القيمة في الجانب الأيسر. طرق توليد تنبؤات السلاسل الزمنية الهرمية:

1. الطريقة التصاعدية **The Bottom-up approach**: تعد من الطرق البسيطة لتوليد تنبؤات متسقة (Athanasopoulos et al., 2009)، حيث يتم توليد تنبؤات لكل سلسلة في المستوى السفلي، ثم جمعها لإنتاج تنبؤات لكل السلاسل في الهيكل على سبيل المثال، بالنسبة للتسلسل الهرمي في الشكل (1)، نقوم أولاً بإنشاء تنبؤات h -خطوة للأمام لكل سلسلة من المستوى الأدنى فنحصل على:

$$\hat{y}_{AA,h}, \quad \hat{y}_{AB,h}, \quad \hat{y}_{AC,h}, \quad \hat{y}_{BA,h}, \quad \hat{y}_{BB,h}$$

بجمع هذه التنبؤات، نحصل على تنبؤات متسقة h -خطوة للأمام لبقية السلسلة:

$$\tilde{y}_h = \hat{y}_{AA,h} + \hat{y}_{AB,h} + \hat{y}_{AC,h} + \hat{y}_{BA,h} + \hat{y}_{BB,h}$$

$$\tilde{y}_{A,h} = \hat{y}_{AA,h} + \hat{y}_{AB,h} + \hat{y}_{AC,h} \quad \text{و} \quad \tilde{y}_{B,h} = \hat{y}_{BA,h} + \hat{y}_{BB,h}$$

(حيث يستخدم ترميز التلدة للإشارة إلى تنبؤات متسقة). كما يمكن تمثيل طريقة Bottom-up مصفوفياً كما يلي:

$$\tilde{y}_h = S \hat{b}_h$$

حيث \tilde{y}_h عبارة عن متجه بعده n يحوي تنبؤات h -خطوة للأمام متسقة و \hat{b}_h عبارة عن متجه بعده m يتضمن تنبؤات h -خطوة للأمام لكل سلسلة من المستوى السفلي. من مميزات هذا النهج أننا نتنبأ عند المستوى الأدنى للهيكل، وبالتالي لا يتم فقدان أية معلومات بسبب التصنيف.

2. الطريقة التنازلية **Top-down approaches**:

تتضمن هذه الطريقة (Hyndman et al., 2011) أولاً إنشاء تنبؤات للسلسلة الاجمالية y_t ، ثم تصنيفها أسفل التسلسل الهرمي. لتكن p_1, \dots, p_m عبارة عن مجموعة من أوزان التصنيف التي تحدد كيفية توزيع تنبؤات السلسلة الاجمالية للحصول على تنبؤات لكل سلسلة في المستوى الأدنى للهيكل. على سبيل المثال، باستخدام النسب p_1, \dots, p_5 من أجل التسلسل الهرمي في الشكل (1) نحصل على:

$$\tilde{y}_{AA,t} = p_1 \hat{y}_t, \quad \tilde{y}_{AB,t} = p_2 \hat{y}_t, \quad \tilde{y}_{AC,t} = p_3 \hat{y}_t, \quad \tilde{y}_{BA,t} = p_4 \hat{y}_t, \quad y_{BB,t} = p_5 \hat{y}_t$$

بترتيب مجموعة الأوزان في متجه $p = (p_1, \dots, p_m)'$ بعده m ، يمكننا كتابة التمثيل المصفوفي التالي:

$$\tilde{b}_t = p \hat{y}_t$$

بعد إنشاء تنبؤات h -خطوة للأمام، يتم تصنيفها لإنشاء تنبؤات متسقة لبقية السلسلة. وبشكل عام، من أجل مجموعة محددة من الأوزان، يمكن تمثيل طريقة Top-down كما يلي:

$$\tilde{y}_h = S p \hat{y}_t$$

حيث تتحدد أوزان التصنيف في طرق Top-down الأكثر شيوعاً بناءً على إحدى الأوزان التالية:

✚ معدل الأوزان الخام Average Historical proportions:

من أجل $j=1, \dots, m$ فإن كل نسبة p_j هي متوسط الأوزان الخام لسلسلة المستوى الأدنى $y_{j,t}$ خلال الفترة $t=1, \dots, T$ نسبة إلى إجمالي السلسلة y_t .

$$p_j = \frac{1}{T} \sum_{t=1}^T \frac{y_{j,t}}{y_t}$$

✚ أوزان المعدلات الخام Proportions of the Historical averages:

من أجل $j=1, \dots, m$ فإن كل نسبة p_j تعطى وفق العلاقة التالية:

$$p_j = \frac{\sum_{t=1}^T \frac{y_{j,t}}{T}}{\sum_{t=1}^T \frac{y_t}{T}}$$

وهي متوسط القيمة الخام لسلسلة المستوى الأدنى $y_{j,t}$ بالنسبة إلى متوسط قيمة إجمالي السلسلة y_t . إن السمة المميزة لطرق Top-down هي بساطتها حيث يحتاج المرء فقط إلى نموذج وإنشاء تنبؤات لسلسلة المستوى الأعلى الأكثر تجميعاً. وبشكل عام، يمكن لهذه الطرق إنتاج تنبؤات موثوقة تماماً للمستويات المجمعّة وهي مفيدة مع البيانات قليلة العدد. من ناحية أخرى، يتمثل أحد عيوب هذه الطريقة في فقدان المعلومات بسبب التصنيف.

✚ أوزان التنبؤ Forecast proportions:

إن تطبيق طريقة Top-down المعتمدة على الأوزان الخام ينتج تنبؤات أقل دقة عند المستويات الدنيا للتسلسل الهرمي مقارنة بالطريقة Bottom-up. ولمعالجة هذه المشكلة، يمكن استخدام الأوزان المعتمدة على التنبؤات بدلاً من البيانات الخام. فإذا كان لدينا تسلسل هرمي بمستوى واحد، عندئذٍ يتم توليد تنبؤات h -خطوة للأمام لجميع السلاسل ولا تستخدم هذه التنبؤات مباشرة، كونها غير متسقة وتسمى هذه التنبؤات بالأولية Initial. ثم تحسب نسبة كل تنبؤ أولي h -خطوة للأمام في المستوى الأدنى إلى إجمالي جميع التنبؤات الأولية h -خطوة للأمام عند هذا المستوى حيث تدعى بأوزان التنبؤ Forecast Proportions، والتي تستخدم لتصنيف التنبؤ الأولي h -خطوة للأمام للمستوى الأعلى وبالتالي يتم توليد تنبؤات متسقة لكامل التسلسل الهرمي.

أما بالنسبة للتسلسل الهرمي بـ K -مستوى، فيتم تكرار هذه العملية من أجل كل عقدة من الأعلى إلى المستوى السفلي. وبتطبيق هذه العملية نحصل على القاعدة العامة التالية للحصول على أوزان التنبؤات:

$$p_j = \prod_{\ell=0}^{K-1} \frac{\hat{y}_{j,h}^{(\ell)}}{\hat{s}_{j,h}^{(\ell+1)}}$$

حيث $\hat{y}_{j,h}^{(\ell)}$ هي التنبؤ الأولي h -خطوة للأمام لسلسلة الموافقة للعقدة الواقعة أعلى العقدة z بـ ℓ مستوى و $\hat{s}_{j,h}^{(\ell)}$ هو مجموع التنبؤات الأولية h -خطوة للأمام لأسفل العقدة الواقعة أعلى العقدة z بـ ℓ مستوى وتتصل مباشرة بتلك العقدة.

تقوم أوزان التنبؤات هذه بتصنيف التنبؤات الأولية h -خطوة للأمام لسلسلة Total للحصول على تنبؤات متسقة h -خطوة للأمام لسلسلة المستوى الأدنى. ومن عيوب كل طرق Top-down بأنها لا تنتج تنبؤات متسقة غير متحيزة.

3. الطريقة الوسطى Middle-out approach:

تجمع طريقة Middle-out بين طريقتي Bottom-up و Top-down (Hyndman et al., 2014)، حيث يتم أولاً اختيار المستوى المتوسط "Middle level" ويتم إنشاء التنبؤات لجميع السلاسل عند هذا المستوى. بالنسبة للسلسلة فوق المستوى المتوسط، يتم إنشاء تنبؤات متسقة باستخدام الطريقة Bottom-up من خلال تجميع تنبؤات المستوى المتوسط إلى الأعلى. بالنسبة للسلسلة الموجودة أسفل المستوى المتوسط، يتم إنشاء تنبؤات متسقة باستخدام طريقة Top-down عن طريق تصنيف تنبؤات المستوى المتوسط إلى أسفل نزولاً.

4. طريقة التجميع الأمثل Optimal Combination :

تعتمد هذه الطريقة على التنبؤ بشكل مستقل بكل السلاسل في جميع مستويات التسلسل الهرمي (Hyndman et al., 2016) ثم استخدام نموذج الانحدار لتركيب وتوفيق هذه التنبؤات على النحو الأمثل، تضاف التنبؤات المنقحة الناتجة بشكل مناسب عبر التسلسل الهرمي، وهي غير منحاذاة ولها تباين أصغري بين جميع التنبؤات المركبة في ظل بعض الافتراضات البسيطة. ومن أساليبها:

- التنبؤ الهرمي العام:

تدعى التنبؤات بالتنبؤات الأساسية "Base forecasts" تلك التي نقوم بحسابها بشكل مستقل لكل سلسلة عند كل مستوى في التسلسل الهرمي باستخدام تاريخ السلسلة فقط (Hyndman and George, 2018) دون أخذ بالحسبان أي علاقة مع سلسلة أخرى في التسلسل الهرمي. إن مجموع هذه التنبؤات الأساسية عند المستويات الدنيا من غير المحتمل أن يكون مساوي للتنبؤات الأساسية لعقد آباءها (سلاسل المستوى الأعلى). وهذا غير متوافق مع بنية تركيب التسلسل الهرمي. وحتى تصبح هذه التنبؤات الأساسية متسقة مع هيكل التسلسل الهرمي، يترتب علينا تعديلها إلى تنبؤات معدلة مجمعة لتكون مساوية للتنبؤات المعدلة للعقدة الأبوية. ونشير إلى التنبؤات المعدلة بالتنبؤات الموقفة Reconciled forecasts.

بفرض تم أولاً حساب التنبؤات لكل سلسلة عند كل مستوى مع إعطاء m تنبؤاً أساسياً لكل من الفترات $n+1, \dots, n+h$ استناداً إلى المعلومات المتاحة المتضمنة حتى اللحظة n . نشير إلى هذه التنبؤات الأساسية بواسطة $\hat{Y}_{X,n}(h)$ حيث تشير X إلى السلسلة التي يتم التنبؤ بها و $\hat{Y}_n(h)$ تشير إلى التنبؤات الأساسية $-h$ خطوة للأمام للمجموع Total.

ليكن $\hat{Y}_n(h)$ متجهاً يتكون من هذه التنبؤات الأساسية، مرتبة في نفس ترتيب السلسلة كما في Y_t . يمكن بعد ذلك كتابة جميع طرق التنبؤ الهرمية الموجودة على النحو التالي:

$$\tilde{Y}_n(h) = SP\hat{Y}_n \quad (4)$$

حيث P مصفوفة مختارة بشكل مناسب من المرتبة $m \times n$. أي أن الطرق الحالية تتضمن تركيبات خطية بالتنبؤات الأساسية. ثم يتم توفيق هذه التركيبات الخطية بمعنى أن تنبؤات المستوى الأدنى تجمع لإعطاء تنبؤات مستوى أعلى. يتمثل تأثير المصفوفة P في استخلاص العناصر ذات الصلة بالتنبؤات الأساسية $\hat{Y}_n(h)$ والتي يتم جمعها بعد ذلك بواسطة S لإعطاء التنبؤات الهرمية المنقحة النهائية $\tilde{Y}_n(h)$ على سبيل المثال، يتم الحصول على تنبؤات Bottom-up باستخدام:

$$P = \left[0_{m_K \times (m-m_K)} \mid I_{m_K} \right] \quad (5)$$

حيث 0 هي مصفوفة صفرية و I هي مصفوفة واحدة. في هذه الحالة، تستخرج المصفوفة P فقط تنبؤات المستوى الأدنى من $\hat{Y}_n(h)$ والتي يتم جمعها بعد ذلك بواسطة S لإعطاء التنبؤات Bottom-up. بينما يتم الحصول على تنبؤات Top-down باستخدام:

$$P = \begin{bmatrix} p & 0_{m_K \times (m-1)} \end{bmatrix} \quad (6)$$

حيث $p = [p_1, p_2, \dots, p_{m_K}]^t$ هو متجه الأوزان. إن تأثير المصفوفة P هنا هو تصنيف تنبؤات الإجمالي إلى سلاسل المستوى الأدنى. لكل من طرق Top-down للتنبؤ المختلفة متجهات P متناسبة مختلفة. كما أن التنبؤات Middle-out ممكنة عن طريق تحديد المصفوفة P بشكل مناسب. ما سبق يشير إلى أنه يمكن تحديد طرق التنبؤ الهرمية عن طريق اختيار مصفوفة P مختلفة بشرط وضع بعض القيود على P لتقديم تنبؤات منطقية.

$$E[\hat{Y}_n(h)] = E[Y_n(h)]$$

وحتى تكون التنبؤات الهرمية المنقحة غير متحيزة، يجب أن يتحقق:

$$E[\tilde{Y}_n(h)] = E[Y_n(h)] = SE[Y_{K,n}(h)]$$

بفرض أن $\beta_n(h) = E[Y_{K,n+h} | Y_1, \dots, Y_n]$ هو متوسط القيم المستقبلية للمستوى الأدنى K عندئذ:

$$E[\tilde{Y}_n(h)] = SPE[\hat{Y}_n(h)] = SPS\beta_n(h)$$

لذا، فإن شرط عدم التحيز للتنبؤات المنقحة هو:

$$SPS = S \quad (7)$$

يتحقق هذا الشرط بالنسبة للطريقة Bottom-up من أجل P معطاة بواسطة (5). إلا أن استخدام الطريقة Top-down مع P المعطاة بواسطة (6) فإن $SPS \neq S$ لذا من أجل أي اختيار لـ P لذا فإن الطريقة Top-down لا يمكنها أبداً إعطاء تنبؤات غير متحيزة حتى لو كانت التنبؤات الأساسية غير متحيزة.

ليكن التباين في التنبؤات الأساسية $\hat{Y}_n(h)$ معطى بواسطة Σ_h عندئذ فإن تباين التنبؤات المنقحة يعطى وفق العلاقة:

$$Var[\tilde{Y}_n(h)] = SP\Sigma_h P^t S \quad (8)$$

وبالتالي، يمكن الحصول على فترات التنبؤ على التنبؤات المنقحة بشرط أن يتم تقدير Σ_h بشكل موثوق .

- التنبؤات المثلى باستخدام الانحدار:

اقترح Hyndman كتابة التنبؤات الأساسية بشكل نموذج خطي كما يلي:

$$\hat{Y}_n(h) = S\beta_n(h) + \varepsilon_h \quad (9)$$

حيث $\beta_n(h) = E[Y_{K,n+h} | Y_1, \dots, Y_n]$ هو متوسط القيم المستقبلية المجهول للمستوى الأدنى K، و ε_h له متوسط صفري ومصفوفة التباين $\Sigma_h = Var(\varepsilon_h)$.

يمكننا تقدير $\beta_n(h)$ باعتبار (9) كمعادلة انحدار، وبالتالي الحصول على تنبؤات لجميع مستويات التسلسل الهرمي. إذا كان Σ_h معلوماً، عندئذ يمكننا استخدام تقدير المربعات الصغرى المعمم للحصول على مقدر غير متحيز ذو تباين أصغري

$$\hat{\beta}_n(h) = (S^t \Sigma_h^t S)^{-1} S^t \Sigma_h^t \hat{Y}_n(h) \quad (10)$$

كما يلي:

حيث Σ_h^t هو معكوس مور-بنروز المعمم Moore-Penrose inverse لـ Σ_h .

(تم استخدام معكوساً معمماً لأن Σ_h غالباً ما تكون شاذة بسبب التجميع المتضمن في Y_n).

مما يؤدي إلى التنبؤات المنقحة التالية: $\tilde{Y}_n(h) = S\hat{\beta}_n(h) = SP\hat{Y}_n(h)$ حيث: $P = (S' \Sigma_h^t S)^{-1} S' \Sigma_h^t$ من الواضح أن هذا يحقق خاصية عدم التحيز (7). كما يعطى تباين هذه التنبؤات من خلال العلاقة التالية:

$$\text{Var}[\tilde{Y}_n(h)] = S(S' \Sigma_h^t S)^{-1} S'$$

تتمتع صعوبة هذه الطريقة في أنها تتطلب معرفة Σ_h ، أو على الأقل تقديراً جيداً لها. في التسلسل الهرمي الكبير، مع الآلاف من السلاسل، قد لا يكون ذلك ممكناً. ومع ذلك، يمكننا تبسيط العمليات الحسابية بشكل كبير من خلال افتراض أن الخطأ في (9) يمكن التعبير عنه بـ $\varepsilon_{K,h} \approx S \varepsilon_h$ ، حيث $\varepsilon_{K,h}$ هو خطأ التنبؤ في المستوى الأدنى. أي أننا نفترض أن الأخطاء تحقق نفس قيد التجميع مثل البيانات الأصلية. سيكون هذا الافتراض صحيحاً بشرط أن تحقق التنبؤات أيضاً قيود هذا التجميع تقريباً.

معايير قياس دقة التنبؤ:

يوجد العديد من المعايير المختلفة لمقارنة جودة التنبؤ الناتج عن طرق التنبؤ الهرمي وهي متوسط الأخطاء التربيعية MSE أو جذر متوسط الأخطاء التربيعية RMSE، متوسط الأخطاء المطلقة MAE ومتوسط الأخطاء النسبية المطلقة MAPE ومتوسط الخطأ المطلق في القياس MASE وهذه المعايير معرفة بالعلاقات التالية:

$$MSE = \frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_t)^2 \quad MAE = \frac{1}{T} \sum_{t=1}^T |y_t - \hat{y}_t|$$

$$RMSE = \sqrt{MSE} \quad MAPE = \frac{1}{T} \sum_{t=1}^T \frac{|y_t - \hat{y}_t|}{|y_t|} (100\%)$$

$$MASE = \text{mean}(|q_t|), \quad q_t = \frac{e_t}{\frac{1}{n-1} \sum_{i=2}^n |Y_i - Y_{i-1}|}$$

تبين من خلال الدراسات السابقة (Hyndman and Koehler, 2006) ومن خلال دراستنا التجريبية في هذا البحث أن MASE

(Mean Absolute Scaled Error) هو مقياس دقة التنبؤ الوحيد الذي يعطي دائماً نتائج معقولة لمقارنة طرق التنبؤ الأربعة ولجميع أنواع السلاسل، حيث تكون الطريقة أفضل كلما انخفضت قيمة MASE. تم اقتراح هذا المقياس عام 2005 من قبل العالم الإحصائي وبروفيسور علم القرار Anne B. Koehler حيث أظهروا بأنه لا ينتج عند تطبيق مقياس دقة التنبؤ هذا المشاكل التي تظهر في القياسات الأخرى كظهور القيم اللانهائية لذلك يوصى به لتحديد دقة التنبؤ الأفضل.

النتائج والمناقشة:

- تصنيف البيانات: تم إنشاء سلسلة زمنية هرمية لعينة بيانات إنتاج القمح الطري لعشر مدن في سوريا (حمص-حماه-الغاب-ادلب-حلب-طرطوس-اللاذقية-الرقعة-دير الزور-الحسكة) والمأخوذة سنوياً خلال الأعوام 2006 - 2018. حيث يشير إجمالي إنتاج القمح الطري في سوريا بأكمله إلى المستوى صفر (المستوى الأعلى) ثم صنفت بيانات إنتاج القمح الطري أولاً حسب المنطقة:

- شرقية (الحسكة، دير الزور، الرقعة)
- غربية (اللاذقية، طرطوس)
- شمالية (حلب، ادلب)
- وسطى (حمص، حماه، ادلب)

حيث تم الحصول على أربع عقد في المستوى الأول للسلسلة، بعد ذلك صنفنا بيانات إنتاج القمح الطري حسب المدينة إلى عشر عقد في المستوى الثاني وهو المستوى الأدنى. كما تم تسمية سلاسل المستوى الأدنى وفق كود مكون من ستة رموز حيث تدل الأحرف الأربعة الأولى من اسم السلسلة على اسم المنطقة (صنف المستوى 1) وآخر رمزين هما عبارة عن رقمين يدلان على المدينة (صنف المستوى 2) حيث تم إعطاء أرقام متسلسلة للمدن ابتداء من الواحد ضمن كل منطقة كما موضح في الجدول التالي:

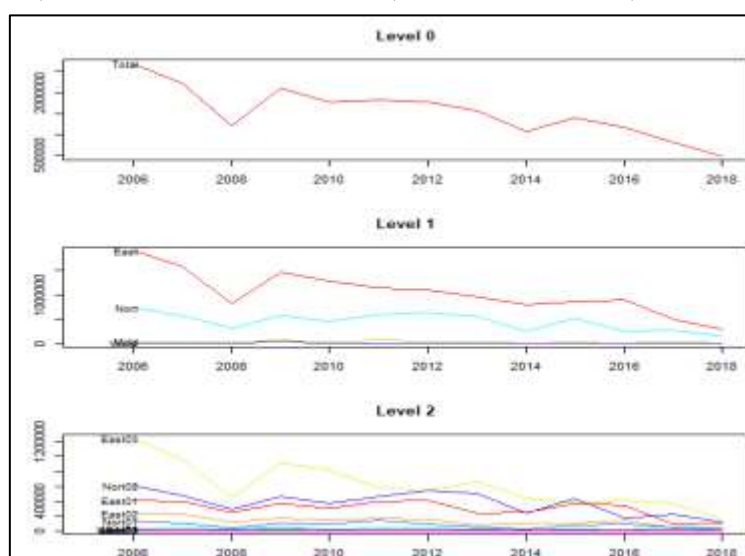
الجدول 1. ترميز أسماء سلاسل التركيب الهرمي لإنتاج القمح الطري في سورية.

الحسكة	دير الزور	الرقبة	اللاذقية	طرطوس	حلب	ادلب	الغاب	حمّاه	حمص
East0	East0	East0	West0	West0	Nort0	Nort0	Midd0	Midd0	Midd0
3	2	1	2	1	2	1	3	2	1

التركيب الهرمي للسلسلة موضح في الشكل (2) حيث يظهر المخطط الأعلى (المستوى 0) إجمالي إنتاج القمح الطري في سوريا، بينما يظهر المخطط الأوسط البيانات المصنفة حسب المنطقة، والتي تكشف عن ديناميكيات متنوعة وغنية على المستوى الإجمالي والمستوى الأول من التصنيف لكل منطقة، أما المخطط الأخير يظهر السلاسل الزمنية للمستوى الأدنى وهو إنتاج القمح الطري حسب المدينة وهذا يساعد على تصور الديناميكيات الفردية المتنوعة داخل كل مدينة، ويساعد في تحديد السلاسل الزمنية الفريدة والمهمة، ومن دراسة الشكل (2) نلاحظ تناقص الإنتاج في إنتاج القمح الطري الإجمالي بينما أظهرت المدن إنتاجاً كبيراً خلال السنوات القليلة الماضية كما في الحسكة وحلب. على سبيل المثال:

في عام (2006) بلغ الإنتاج الكلي للقمح الطري في سورية (2701135 طن) موزع في المناطق الشرقية والغربية والشمالية والوسطى على الترتيب (1892646)، (16657)، (742119)، (49713)، من خلال إنتاج المدن: الحسكة ودير الزور والرقبة على الترتيب (1250872)، (234157)، (407617)، اللاذقية وطرطوس (92) و(16565)، حلب وادلب (609695) و (132424)، حمص وحمّاه وادلب (14159)، (11051)، (24503).

وفي عام (2007) كان الإنتاج الكلي (2224703 طن) وفي عام (2008) تناقص الإنتاج الكلي إلى (1199603 طن).



الشكل (2): السلاسل الزمنية لمستويات التركيب الهرمي لإنتاج القمح الطري في سورية

- الطريقة المقترحة للتنبؤ:

1. تم بناء نموذج هرمي على البيانات من عام 2006 لعام 2018.
2. قسمت بيانات السلسلة الهرمية إلى مجموعة تدريب من عام 2006 حتى عام 2014 ومجموعة اختبار من عام 2015 حتى 2018.
3. تم بناء النموذج الأفضل لكل سلسلة باستخدام طريقة التنعيم الأسّي Exponential Smoothing ثم استخدمت هذه النماذج لحساب التنبؤ وفق الطرق الهرمية المذكورة: التجميع الأمثل Optimal Combination والطرق التقليدية Middle-out و Bottom-up، Top-down.
4. قدر مؤشر متوسط الخطأ المطلق MASE لمجموعة الاختبار من أجل كل سلسلة في التسلسل الهرمي. حيث نجد كما هو وارد في الجدول رقم (2):
 - طريقة Top-down هي الأفضل أداء بشكل عام عبر جميع مستويات التسلسل الهرمي.
 - وفي المستوى الاجمالي تتفوق طريقة Top-down وطريقة Middle-out بشكل واضح على طريقتي Bottom-up و Optimal Combination.
 - وفي المستوى الأول نجد أن دقة التنبؤ للطرق ماعدا Top-down هي نفسها في كل من المنطقة الوسطى والغربية.
 - وفي المستوى الثاني (الأدنى) تتفوق طريقة Middle-out على بقية الطرق بالنسبة لسلسلة بيانات مدينة الحسكة. كما نجد نفس الدقة للطرق باستثناء Top-down من أجل كل من سلسلة بيانات مدينة حمص وحماه والغاب أما بالنسبة لسلسلة بيانات مدينة طرطوس في المستوى الأدنى نجد أنه للطرق الأربعة تقريباً نفس الدقة في التنبؤ. وهكذا بعد مقارنة أداء التنبؤ ببيانات إنتاج القمح الطري في سوريا لكل من طرق التنبؤ الأربعة، اخترنا أفضل طريقة وهي طريقة Top-Down.

الجدول 2. قيم MASE لكل سلاسل التركيب الهرمي

	Top-down	Middle-out	Bottom-up	Optimal Combination
Total	0.441	0.488	0.687	0.578
East	0.506	0.215	0.431	0.309
Midd	0.209	0.677	0.677	0.676
Nort	0.477	0.832	0.976	0.903
West	0.48	0.513	0.504	0.508
East01	0.454	0.55	0.754	0.724
East02	0.675	0.113	0.173	0.149
East03	0.493	0.096	0.237	0.106
Midd01	0.487	0.832	0.833	0.832
Midd02	0.201	0.531	0.531	0.531
Midd03	0.22	0.491	0.491	0.49
Nort01	0.235	0.582	0.675	0.644
Nort02	0.512	0.839	0.985	0.906
West01	0.48	0.51	0.501	0.505
West02	0.126	0.269	0.264	0.265
Average	0.3997	0.503	0.581	0.541

المصدر: من عمل الباحث باستخدام لغة R

5. تم تجربة النموذج وفق الطريقة الأفضل أداء على مجموعة بيانات الاختبار من عام 2015 لعام 2018 فكانت دقة التنبؤ مقبولة كونها أكبر من 80% لكل سلاسل التركيب الهرمي، الجدول (3، 4). ولذلك يمكن الاعتماد على النموذج المقترح في التنبؤ.

الجدول 3. دقة التنبؤ لمجموعة بيانات الاختبار على المستوى الإجمالي والمناطق.

Total	East	Midd	Nort	West	
0.86	0.87	0.84	0.85	0.83	2015
0.87	0.88	0.85	0.85	0.85	2016
0.93	0.94	0.92	0.88	0.81	2017
0.85	0.85	0.85	0.86	0.87	2018

المصدر: من عمل الباحث باستخدام لغة R

الجدول 4. دقة التنبؤ لمجموعة بيانات الاختبار على مستوى المدن.

East01	East02	East03	Midd01	Midd02	Midd03	Nort01	Nort02	West01	West02	
0.99	0.79	0.82	0.82	0.91	0.83	0.91	0.84	0.83	0.88	2015
0.99	0.81	0.85	0.82	0.92	0.84	0.91	0.83	0.84	0.86	2016
0.87	0.89	0.86	0.89	0.96	0.83	0.86	0.89	0.81	0.9	2017
0.8	0.8	0.89	0.87	0.8	0.85	0.9	0.85	0.87	0.86	2018

المصدر: من عمل الباحث باستخدام لغة R

6. التنبؤ: طبقت الطريقة المقترحة TOP_Down للتنبؤ بثلاث سنوات لاحقة بعد عام 2018، وتشير قيم التنبؤ بحصول زيادة في إنتاج محصول القمح الطري على المستوى الإجمالي وجميع مستويات السلسلة الهرمية، الجدول (5، 6). ومن خلال بيانات المجموعة الإحصائية 2019 التي صدرت مؤخراً قد لوحظ فعلاً أن محصول القمح الطري سيتزايد في جميع المستويات وخاصة المستوى الإجمالي الذي بلغ (1321296 طن) وهي قيمة قريبة عن القيمة المتنبأ بها من قبل النموذج المطبق (1208321 طن)، والقيمة المتنبأ لإنتاج مدينة اللاذقية عام 2019 هي (42.3 طن) لا تتعد عن القيمة الحقيقية (32 طن) وهذا يدعم النموذج المقترح واتجاهه في التزايد.

الجدول 5. القيم المتنبأ لسلاسل إنتاج القمح الطري بالطن من عام 2019 حتى 2021 على المستوى الإجمالي والمناطق.

Total	East	Midd	Nort	West	
1208321	812633	31874.3	354208	9605.9	2019
1207830	812303	31861.3	354064	9602	2020
1207859	812322	31862.1	354073	9602.2	2021

المصدر: من عمل الباحث باستخدام لغة R.

الجدول 6. القيم المتنبأ لسلاسل إنتاج القمح الطري بالطن من 2019 حتى 2021 على مستوى المدن.

East01	East02	East03	Midd01	Midd02	Midd03	Nort01	Nort02	West01	West02	
23079	1004	48136	7841.	4944	19089	57692	2965	9563.	42.3	2019
9.4	69	3.9	2			.9	15	6		9

23070	1004	48116	7838	4942	19081	57669	2963	9559.	42.2	202
5.7	29	8.5			.2	.5	95	7		0
23071	1004	48117	7838.	4942	19081	57670	2964	9560	42.2	202
1.1	31	9.8	2		.7	.9	02			1

المصدر: من عمل الباحث باستخدام لغة R.

الاستنتاجات والتوصيات:

تم في هذا البحث مناقشة طرق التنبؤ بالسلسلة الزمنية الهرمية: Top-Down، Bottom-Up، و Middle-Out مع الطريقة الإحصائية الجديدة للتنبؤ Optimal Combination، إحدى الميزات المفيدة لتلك الطريقة هي أن التنبؤات الأساسية يمكن أن تأتي من أي نموذج حيث أنها لا تضع قيوداً على كيفية إنشاء التنبؤات الأساسية. كما توفر وسيلة لتوفيق التنبؤات بحيث يتم تجميعها بشكل مناسب عبر التسلسل الهرمي. ميزة ملحوظة أخرى هي أن التنبؤات النقطية مستقلة عن الارتباط بين السلاسل. في حين أن هذا قد يبدو مبدئياً غير بديهي، إلا أنه نتيجة طبيعية لافتراض أن أخطاء التنبؤ عبر التسلسل الهرمي تتجمع بنفس الطريقة التي يتم بها تجميع البيانات المشاهدة، وهو تقريب معقول للواقع. إلا أن تباينات التنبؤ تعتمد بشكل حاسم على الارتباط بين السلاسل، حيث أن إنتاج فترات التنبؤ يتطلب مصفوفة التغيرات. نتيجة أخرى هي أن أوزان التجميع الأمثل تعتمد فقط على الهيكل الهرمي وليس على البيانات المشاهدة. ينشأ هذا أيضاً من الافتراض المتعلق بتجميع أخطاء التنبؤ. بسبب هذه النتيجة، يمكن تحديد أوزان التجميع مرة واحدة فقط، ثم تطبيقها عند توفر كل مجموعة جديدة من المشاهدات مما يوفر الكثير من الوقت الحسابي.

وأنت النتائج بتطبيق هذه المنهجيات على إنتاج القمح الطري في سورية على المستوى الإجمالي والمناطق والمدن لسلسلة هرمية من عام 2006 إلى عام 2018، وتم مقارنة النتائج التي تم الحصول عليها وفق الطرق المذكورة واختبار النتائج التجريبية فتوضح أن طريقة Top-down تتفوق على الطرق الموجودة للتنبؤ بالبيانات الهرمية لإنتاج القمح الطري في سورية حسب قيم معيار MASE. وتم التنبؤ لثلاث سنوات مستقبلية بعد عام 2018 لتأتي النتائج بتوقع زيادة في إنتاج القمح الطري في جميع مستويات التركيب الهرمي، لذلك يوصى باعتمادها للاستخدام الروتيني في الأعمال الزراعية، كما كانت هناك حاجة للتنبؤ ببيانات السلاسل الزمنية الهرمية.

المراجع:

جابر، علي حسين ومجيد وأروى مهند (2017). تحليل درجات الحرارة في مدينة بغداد للفترة (2015-2016) باستخدام السلاسل الزمنية. بحث بكالوريوس، كلية علوم الحاسوب وتكنولوجيا المعلومات، جامعة القادسية، العراق. 49 صفحة.
وزارة الزراعة والإصلاح الزراعي (2006-2018). المجموعة الإحصائية الزراعية. قسم الإحصاء، مديرية التخطيط والتعاون الدولي، دمشق، سورية.

Athanasopoulos, G.; R. A. Ahmed; and R.J. Hyndman (2009). Hierarchical forecasts for Australian Domestic tourism. *International Journal of Forecasting*. 25(1): 146-166.

Hyndman, R. J.; George Athanasopoulos; and Han Lin Shang (2014). hts: An R Package for Forecasting Hierarchical or Grouped Time Series. Monash University.

Hyndman, R. J.; and A.B. Koehler (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*. Monash University. 22(4): 679-688.

Hyndman, R. J.; A. Lee; and E. Wang (2016). Fast computation of reconciled forecasts for hierarchical and grouped time series. *Computational Statistics and Data Analysis*. 97: 16-32.

- Hyndman, R. J.; Yeasmin Khandakar (June 2007), Automatic time series forecasting: the forecast package for R. Monash University. VIC 3800. Australia. ISSN 1440-771X.
- Hyndman, R. J.; and George Athanasopoulos (April 2018). Forecasting: Principles and Practice, Monash University. Australia.
- Hyndman, R. J.; R. A. Ahmed; G. Athanasopoulos and H. L. Shang (2011). Optimal combination forecasts for hierarchical time series, *Computational Statistics and Data Analysis*.55(9): 2579–2589.
- Simon H. Heisterkamp; L. M. Arnold Dekkers; and C. M. Janneke Heijne (2006). Automated Detection of infectious disease outbreaks: hierarchical time series models. *Statistics in Medicine*. 25:4179–4196.
- Wickramasuriya, S. L.; G. Athanasopoulos and R. J. Hyndman (2019). Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization. *J American Statistical Association*. 114(526): 804–819.

Comparing The Quality of Some Short-Term Forecasting Methods of The Hierarchical Time Series Data for Bread Wheat Production in Syria

Falak Alsataihi *⁽¹⁾, Mohamad Taher Anan ⁽²⁾ and Amjad Masso⁽³⁾

(1). Department of Basic Sciences, Faculty of Chemical and Petroleum Engineering, AlBaath University, Homs, Syria.

(2). Department of mathematical Statistics and Programming, Faculty of Science, Aleppo University, Aleppo, Syria.

(*Corresponding author: Falak Alsataihi, E-Mail: falakalsataihi@yahoo.com).

Received: 31/03/2021

Accepted: 4/07/2021

Abstract

The research studied the methods of generating short-term hierarchical time-series forecasts and applying them to soft wheat production data in Syria, which were classified by region and then by city for time-series from 2006-2018. Where four methods were discussed, namely Top-down, Bottom-up, Middle-out and Optimal Combination, and one of the criteria for testing the quality of prediction methods, which is MASE, was applied to find out the best-performing method to be adopted in research related to this research in order to facilitate and speed up the completion of calculations. The results showed that all four methods differ from each other in general. As a result of the pilot study, it was found that the performance of the Top-down method is better than the performance of the rest of the methods in forecasting, and it was predicted for three future years after 2018, where the results indicated an expectation of an increase in the production of the soft wheat crop at all levels of the hierarchical structure of the studied series.

Keywords: Hierarchical Time Series, Bottom-up approach, Top-down approaches, Middle-out approach, Optimal combination forecasts.