

نموذج إحصائي للتنبؤ بكميات إنتاج وإستهلاك المياه النقية في محطات المياه التابعة للشركة القابضة في مصر باستخدام أسلوب الدمج بين نموذجي الإنحدار النشط و إنحدار ريديج

د/ هناء حسين على أبو العلا
مدرس بقسم الإحصاء والرياضة والتأمين
كلية التجارة - جامعة عين شمس

مستخلص

يهدف البحث إلى معالجة مشكلتي الإزدواج الخطي والقيم الشاذة باستخدام أسلوب الدمج بين نموذجي الإنحدار النشط وإنحدار ريديج على محطات المياه النقية التابعة للشركة القابضة للمياه في مصر هذا بالإضافة للتنبؤ بإنتاج وإستهلاك المياه في هذه المحطات وذلك من عام 2020 الى عام 2030 ، وتمت المقارنة بين كل من نموذج الإنحدار المتعدد بطريقة المربعات الصغرى (Least Squares (OLS) ، ونموذج إنحدار Ordinary Regression، ونموذج الإنحدار النشط Robust Regression ، ونموذج إنحدار ريديج Ridge Regression بالإضافة لأسلوب الدمج بينهما Robust Ridge Regression من حيث القدرة التفسيرية والتنبؤية، وتوصلت النتائج إلى أن أفضل نموذج للتنبؤ بكميات المياه المستهلكة في الشركة القابضة هو نموذج Robust Ridge M Regression حيث إرتفعت قدرته التفسيرية والتنبؤية عن النماذج الأخرى، ووجد أن أهم العوامل المؤثرة على إجمالي كمية المياه المستهلكة (WC_{it}) هي كميات المياه المستهلكة في الوحدات السكنية (HU_{it})، والمحال التجارية (CS_{it}) والورش والمصانع (WF_{it}) ، وشركات السياحة (TC_{it})، وأماكن أخرى (OP_{it}) هذا بالإضافة إلى فاقد الشبكات (NL_{it}) كما تبين أن

أفضل نموذج للتنبؤ بكميات المياه المنتجة في الشركة القابضة هو نموذج **Robust Ridge MM Regression** حيث إرتفعت قدرته التفسيرية والتنبؤية أيضا ووجد أن أهم العوامل المؤثرة على كمية المياه المنتجة (WP_{it}) بالمتر المكعب في محطات المياه التابعة للشركة القابضة هي عدد ساعات التشغيل يوميا (OH) ، وعدد أيام التشغيل سنويا (OD_{it})، والقدرة الفعلية للمحطة (AC_{it}) ، والقدرة التصميمية للمحطة (DC_{it})، وعدد محطات المياه (NS_{it}) .

الكلمات المفتاحية : الموارد المائية ، إنتاج المياه ، إستهلاك المياه ، الإزدواج الخطي ، القيم الشاذة ، نموذج إنحدار المربعات الصغرى العادية ، نموذج الإنحدار النشط ، نموذج إنحدار ريدج ، نموذج إنحدار ريدج النشط ، السلاسل الزمنية.

Abstract

The research aims to solve multicollinearity and outliers problems using the hybrid approach of the two models of robust regression and ridge regression on pure water stations of the Holding Company for Water in Egypt in addition to forecasting the production and consumption of water in these stations from 2020 to 2030. The multiple regression model using the ordinary least squares (OLS) method was compared with robust regression model , ridge regression model as well as the hybrid approach of robust and ridge regression by using predictability and explanatory ability. the statistical results indicated that the best model for predicting of the quantities of water consumed in the holding company is robust ridge M regression model, where its explanatory and predictive ability is higher than other models. Also it was found that the most important factors affecting of the total amount of water consumed (WC_{it}) are the quantities of water consumed in housing units (HU_{it}), commercial shops (CS_{it}), workshops and factories (WF_{it}), tourism companies (TC_{it}), and other Places (OP_{it}) in addition to networks lost (NL_{it}). It was also found that the best model for predicting by the quantities of water produced in the Holding Company is robust ridge MM regression model, where its explanatory and predictive ability also

increased, and it was found that the most important factors affecting of the amount of water produced (WP_{it}) per cubic meter in the Holding Company water stations are the number of operating hours (OH_{it}) daily , number of operating days (OD_{it}) per year , actual capacity of station (AC_{it}), design capacity of station (DC_{it}), number of water stations (NS_{it}).

Keywords :- Water resources , Water pproduction ,Water consumption, Multicollinearity, Outliers , Ordinary least squares regression model , Robust regression model, Ridge regression model, Robust ridge regression, Time series.

أولاً : مقدمة الدراسة :-

إن الماء هو أساس الحياة وهو أهم مدخلات التنمية الإجتماعية والإقتصادية فى العالم وأحد أهم عناصر الإستدامة البيئية ، وتعد المنطقة العربية أحد المناطق الأكثر فقراً فى العالم فيما يتعلق بالموارد المائية ، وتسبب ندرة المياه فى معظم دول المنطقة العربية تحديات كبيرة لجهود التنمية المستدامة على المستوى الوطنى والإقليمى ، وتساهم الزيادة السكانية المطردة والآثار المتوقعة للتغيرات المناخية على المنطقة فى زيادة الضغوط على الموارد المائية المحدودة ، وقد تبنى المجتمع الدولى مبادئ الإدارة المتكاملة للموارد المائية لمواجهة التحديات المتنامية فى قطاع المياه على مستوى العالم وذلك فى كافة الجهود لتحقيق التنمية المستدامة وأخرها أجندة التنمية المستدامة حتى عام 2030 والتي تم إقرارها من الجمعية العامة للأمم المتحدة فى سبتمبر 2015 ، وتهدف الإدارة المتكاملة للموارد المائية بشكل عام إلى تحقيق الكفاءة الإقتصادية والعدالة الإجتماعية والإستدامة البيئية فيما يتعلق بقطاع المياه، وتفتقر الدول العربية بصفة عامة ومصر بصفة خاصة إلى وجود إستراتيجيات متكاملة للموارد المائية تنسق الجهود بين كافة مؤسسات الدولة على المستوى الوطنى لتحقيق التنمية المستدامة (المجلس العربى للمياه ، 2015) [2] .

ثانيا: مشكلة الدراسة :-

إن عدد سكان مصر في تزايد مستمر بينما الموارد المائية من مياه النيل محدودة وهي حوالي 55.5 مليار متر مكعب سنويا منذ إتفاقية السودان مما أدى ذلك إلى تناقص نصيب الفرد من المياه من حوالي 2000 متر مكعب لكل فرد من عام 1959 إلى حوالي 630 متر مكعب لكل فرد عام 2015 ، ومن المتوقع أن يصل عدد سكان مصر عام 2050 إلى حوالي 150 مليون نسمة ممايعنى أن نصيب الفرد من المياه سيقبل إلى حوالي 370 متر مكعب لكل فرد سنويا ، ونلاحظ أن 95% من السكان يعيش على حوالي 5% من المساحة الإجمالية لمصر ويتركز السكان في الدلتا والشريط الضيق حول وادي نهر النيل مما يؤدي إلى كثافات سكانية عالية في تلك المناطق ، وبالتالي كان التفكير في إنشاء عدد من المشروعات القومية التي تعمل على إنشاء تجمعات سكانية جديدة تعمل على توزيع أفضل للسكان وخلق فرص عمل جديدة في كافة المجالات الزراعية والصناعية والسياحية والخدمية أصبح أمر هام وذو ضرورة ملحة ، وتعتمد مصر على مياه النيل كمورد رئيسي للمياه والذي يمثل 97 % من الموارد المائية المتجددة في مصر (وزارة الموارد المائية والرى ، 2017) [9] . وطبقا لإحصائيات الجهاز المركزى للتعبئة العامة والإحصاء ووزارة الموارد المائية والرى بلغ إجمالى الموارد المائية المتاحة حوالي 80.25 مليار متر مكعب لعام 2016 / 2017 ، وذلك نتيجة التوسع فى إستخدام المياه الجوفية ، وإعادة تدوير المياه مرة أخرى وإعادة إستخدامها ، وفيما يلى عرض للموارد والإستخدامات المتاحة للمياه فى مصر، ونبدأ بعرض هذه الموارد المائية المتاحة من خلال الجدول التالى :-

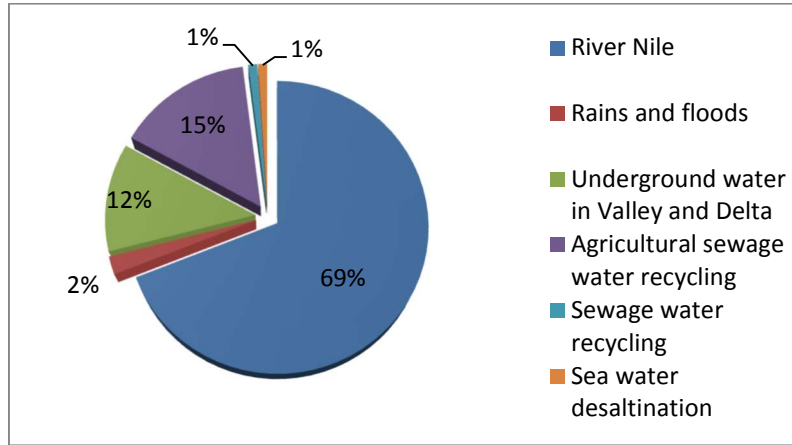
جدول (1) الموارد المائية المتاحة في مصر طبقا للمصدر

(2017-2012) بالمليار متر مكعب سنويا *

| Soure | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 |
|--|--------------|--------------|-------------|--------------|--------------|--------------|
| Share of River Nile (SRN) | 55.50 | 55.50 | 55.50 | 55.50 | 55.50 | 55.50 |
| Rains and floods (RF) | 0.63 | 0.74 | 0.90 | 0.90 | 0.65 | 1.30 |
| Underground water in Valley and Delta (UWVD) | 7.50 | 7.70 | 6.70 | 6.90 | 6.90 | 9.60 |
| Agricultural sewage water recycling (ASWR) | 9.17 | 10.10 | 11.50 | 11.70 | 11.90 | 12.3 |
| Sewage water recycling (SWR) | 1.30 | 1.30 | 1.30 | 1.30 | 1.20 | 1.20 |
| Sea water desalination (SWD) | 0.06 | 0.06 | 0.10 | 0.10 | 0.10 | 0.35 |
| Total | 74.20 | 75.40 | 76.0 | 76.40 | 76.30 | 80.25 |

*المصدر: الجهاز المركزي للتعبئة العامة والإحصاء ، وزارة الموارد المائية والرى.
يتضح من الجدول السابق أن حصة مياه نهر النيل (SRN) بلغت 55.5 مليار متر مكعب وهى لم تتغير منذ إتفاقية السودان عام 1959 حتى الآن كما ذكرنا من قبل وتبلغ حوالى 69.2 % من إجمالى الموارد المائية المتاحة في مصر عام 2017، ويأتى بعد ذلك الموارد المائية غير التقليدية وأهم مواردها تدوير مياه الصرف الزراعى (ASWR) الذى زادت نسبته من 9.17 مليار متر مكعب عام 2012 حتى وصلت إلى 12.3 مليار متر مكعب عام 2017 وهو مايعادل 15 % من إجمالى الموارد المائية المتاحة ، وأما عن المياه الجوفية من الوادى والدلتا (UWVD) سواء الضحلة أو العميقة فهى تعد المورد الثالث فقد بلغت حصتها 7.50 مليار متر مكعب عام 2012 ثم زادت حتى وصلت إلى 9.6 مليار متر مكعب عام 2017 وهو مايعادل 12 % تقريبا من إجمالى الموارد المائية المتاحة عام

2017 ، وأما عن كمية الأمطار والسيول (RF) وهى مصدرا محدودا ولايعتمد عليه فقد زادت من 0.63 مليار متر مكعب عام 2012 حتى وصلت إلى 1.30 مليار متر مكعب عام 2017 وهو مايعادل نسبة 2 % تقريبا من إجمالي الموارد المائية المتاحة عام 2017 ويأتى بعده تدوير مياه الصرف الصحى (SWR) وكانت نسبته 1.3 مليارمتر مكعب عام 2012 ثم إنخفضت إنخفاضا طفيفا حتى وصلت إلى 1.2 مليار متر مكعب عام 2017 وهو مايعادل نسبة 1 % من إجمالي الموارد المائية المتاحة عام 2017 وأخيرا جاء مورد تحلية مياه البحر (SWD) فى المرتبة الأخيرة حيث بلغ 0.06 مليار مترمكعب عام 2012 ثم زاد زيادة طفيفة حتى وصل إلى 0.35 عام 2017 وهو مايعادل 1% من إجمالي الموارد المائية المتاحة عام 2017 ، ويوضح الشكل البيانى التالى نسب الموارد المائية المتاحة فى مصر كمايلى :-



شكل (1) التوزيع النسبى للموارد المائية المتاحة فى مصر عام 2017 وأما عن الإستخدامات المائية المتاحة فى مصر فى الفترة من 2012 حتى 2017 تظهر من الجدول رقم (2) كمايلى:-

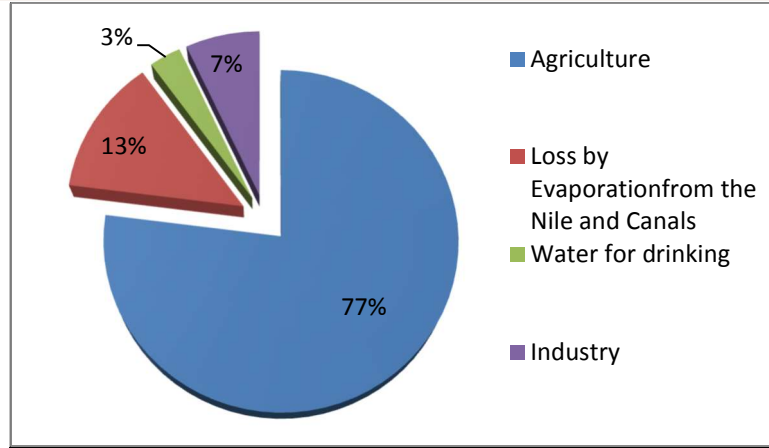
جدول (2) الإستخدامات المائية فى مصر طبقا للمصدر

(2012-2017) بالمليار متر مكعب سنويا *

| Source | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 |
|---|--------------|--------------|--------------|-------------|--------------|--------------|
| Agriculture (Agr.) | 61.10 | 62.10 | 62.35 | 62.35 | 62.15 | 61.65 |
| Loss by Evaporation from the Nile (NC) and Canals | 2.20 | 2.50 | 2.50 | 2.50 | 2.50 | 2.50 |
| Water for (WD) drinking | 9.60 | 9.70 | 9.95 | 10.35 | 10.45 | 10.70 |
| Industry (Indus.) | 1.20 | 1.20 | 1.20 | 1.20 | 1.20 | 5.40 |
| Total | 74.10 | 74.10 | 75.50 | 76.0 | 76.30 | 80.25 |

*المصدر: الجهاز المركزى للتعبئة العامة والإحصاء ، وزارة الموارد المائية والرى .

يتضح من الجدول السابق أن الزراعة هى المستهلك الرئيسى للمياه فى مصر (Agr.) حيث بلغت 61.1 مليار متر مكعب عام 2012 ثم زادت حتى وصلت إلى 61.6 مليار متر مكعب عام 2017 وهو مايعادل 77 % من إجمالى الإستخدامات المائية المتاحة فى مصر عام 2017 ، وأما عن كمية مياه الشرب (WD) فتأتى فى المرتبة الثانية حيث بلغت 9.6 مليار متر مكعب عام 2012 ثم زادت حتى وصلت إلى 10.70 مليار متر مكعب عام 2017 وهو مايعادل 13% من إجمالى الإستخدامات المائية المتاحة فى مصر عام 2017 ثم جاءت كمية المياه المستهلكة فى قطاع الصناعة فى المرتبة الثالثة (Indus.) حيث بلغت حصتها 1.20 مليار متر مكعب عام 2012 ثم زادت حتى وصلت إلى 5.40 مليار متر مكعب عام 2017 وهو مايعادل 7 % من إجمالى الإستخدامات المائية المتاحة عام 2017 وأخيرا جاءت كمية التبخر من النيل والترع (NC) فى المرتبة الرابعة فقد زادت من 2.2 مليار متر مكعب عام 2012 حتى وصلت إلى 2.5 مليار متر مكعب عام 2017 وهو مايعادل نسبة 3% من إجمالى الإستخدامات المائية المتاحة عام 2017، ويوضح الشكل البيانى نسب الإستخدامات المائية المتاحة فى مصر كمايلى :-



شكل (2) التوزيع النسبي للإستخدامات المائية المتاحة في مصر عام 2017 يتضح من العرض السابق أن تنمية الموارد المائية وترشيد إستهلاكها أصبح ضرورة حتمية لتلبية إحتياجات مصر الحالية والمستقبلية. لذلك تهدف الدراسة الحالية إلى بناء النماذج الإحصائية اللازمة لتقدير العوامل المؤثرة على دوال إنتاج وإستهلاك المياه في محطات المياه النقية التابعة للشركة القابضة للمياه في مصر وذلك في الفترة من عام 2000 حتى 2019 ، ثم التنبؤ بدوال الإنتاج والإستهلاك في الفترة من 2020 حتى 2030 وستقتصر الدراسة على محطات المياه النقية للشركة القابضة للمياه نظرا لتوافر بيانات عنها في الفترة المحددة للدراسة ومع العلم أن هناك أنواع أخرى من محطات المياه في مصر منها مايتبع هيئة قناة السويس وأخرى تتبع هيئة المجتمعات العمرانية الجديدة و ثالثة تتمثل في محطات مياه الصرف الصحي التابعة للشركة القابضة للمياه وهذه المحطات لن يتم إستخدامها في الدراسة الحالية لعدم توافر بيانات كاملة عنها ، ثم يقوم الباحث بعد ذلك بإجراء المقارنة بين نموذج إحدار المربعات الصغرى العادية (OLS) ونموذج الإحدار النشط **Robust Regression**، ونموذج إحدار ريدج **Ridge Regression** بالإضافة لأسلوب الدمج **Robust Ridge Regression** ومعرفة أيهم يملك أعلى قدرة تفسيرية

- وتنبؤية لدوال إنتاج وإستهلاك المياه فى الشركة القابضة فى الفترة من 2020 وحتى 2030، وتهدف الدراسة أيضا لإختبار مدى صحة الفروض التالية :-
- (1) لا توجد علاقة ذات دلالة إحصائية معنوية بين إجمالى كمية إستهلاك المياه فى الشركة القابضة (WC_{it}) وبين كل من كمية المياه الموزعة على كل مايلى : (الوحدات السكنية (HU_{it}) ، المحال التجارية (CS_{it}) ، الورش والمصانع (WF_{it}) ، وشركات السياحة (TC_{it}) ، وأماكن أخرى (OP_{it}) ، فاقد الشبكات (NL_{it}) كمغيرات مستقلة .
- (2) لا توجد علاقة ذات دلالة إحصائية معنوية بين إجمالى كمية المياه المنتجة (WP_{it}) وبين كل من (عدد ساعات التشغيل (OH_{it}) ، القدرة الفعلية للمحطة (AC_{it}) ، القدرة التصميمية (DC_{it}) ، عدد المحطات (NS_{it}) كمغيرات مستقلة .
- (3) إن نموذج الدمج **Robust Ridge Regression** أفضل للتنبؤ بكميات المياه المنتجة والمستهلكة فى محطات المياه النقية فى الشركة القابضة من نموذج إحدار المربعات الصغرى العادية (OLS) ، ونموذج الإحدار النشط **Robust Regression**، ونموذج إحدار ريدج **Ridge Regression** .
- (4) يتوقع حدوث زيادة فى إستهلاك محطات المياه النقية فى الشركة القابضة للمياه فى مصر وإنخفاض المعروض منها خلال فترة التنبؤ .
- سوف تعتمد الدراسة فى الحصول على بياناتها من خلال المجالات الإقتصادية والنشرات الإحصائية والتقارير التى تصدرها الجهات التالية (الجهاز المركزى للتعبئة العامة والإحصاء ، وزارة الموارد المائية والرى) .

ثالثاً : الدراسات السابقة :-

تتناول الدراسات السابقة أهم الدراسات النظرية والتطبيقية التي طبقت على نموذج إنحدار ريدج Ridge Regression ونموذج الإنحدار النشط Robust Regression والنموذج الناتج عن الدمج بينهما Robust Ridge Regression وذلك بغرض حل مشكلتي الإزدواج الخطي والقيم الشاذة كمايلي :-

(1) تعتبر دراسة (Holland, 1973) [19] هي أول دراسة قدمت مدخلا لتقدير دالة الإنحدار الخطي المتعدد في ظل وجود مشكلتي الإزدواج الخطي والقيم الشاذة من خلال إستخدام إنحدار ريدج المرجح بأوزان نموذج إنحدار M النشط ، وتوصلت الدراسة إلى قدرة هذا النموذج على حل مشكلتي الإزدواج الخطي والقيم الشاذة معا من خلال خفض معامل تضخم التباين (VIF) Variance Inflation factor لمعاملات الإنحدار المقدره والحصول على مقدرات غير متحيزة .

(2) إقترحت دراسة (Hoerl et al., 1975) [18] بأخذ الوسط التوافقي Harmonic Mean لمعلمة ريدج المعممة وهو (\hat{K}_{HKB}) ، كما إستخدمت دراسة (Lawless and Wang ,1976) [23] المدخل البيزي Bayesian Approach في تقديم مقدر جديد لمعلمة ريدج وهو (\hat{K}_{LW}) وذلك لحل مشكلة الإزدواج الخطي والحصول على معلمة تحيز جديدة لكي تعطى نتائج أفضل لمقدرات نموذج الإنحدار الخطي .

(3) قامت دراسة (Askin & Montgomery, 1980) [11] بالمقارنة بين عائلة من المقدرات التي تجمع بين مقدرات نموذج إنحدار Robust Regression وكل من مقدرات نموذج إنحدار ريدج Ridge Regression ومقدر Stein ونموذج المكونات الرئيسية Principal Components ، وقد تبين من النتائج أن أفضل نموذج لعلاج مشكلة الإزدواج الخطي في المتغيرات المستقلة هو نموذج Ridge Regression لأنه يحقق إرتفاع في القدرة التفسيرية لنموذج الإنحدار الخطي عن باقي النماذج الأخرى.

- (4) قدمت دراسة (Pfaffenberger & Dielman 1984) [27] مقدرًا جديدًا يجمع بين خصائص مقدر الإنحدار النشط وهو مقدر الإنحرافات المطلقة الصغرى (LAD) (Least Absolute Deviations) مع مقدر ريدج وهو مقدر (\hat{K}_{HB}) Hoerl & Kennard .
- (5) قدمت دراسة (Nomura, 1988) [24] مقدرًا جديدًا لمعلمة إنحدار ريدج المعمم (\hat{K}_{NO}) كما قدمت أيضًا صيغة جديدة للمقدر في إنحدار ريدج العادي وذلك بأخذ الوسط التوافقي للصيغة المعممة (\hat{K}_{NO}) .
- (6) إقترحت دراسة (Kibria , 2003) [22] ثلاث مقدرات جديدة لمقدر ريدج المعمم (\hat{K}_{HB}) والتي قدمها كل من Horel and Kennard وهذه المقدرات تتمثل في أخذ الوسط الحسابي (\hat{K}_{KAM}) والوسط الهندسي (\hat{K}_{KGM}) والوسيط (\hat{K}_{KGM}) لصيغة مقدر ريدج المعمم ، وتوصلت الدراسة إلى أن المقدرين (\hat{K}_{KAM}) ، (\hat{K}_{KGM}) كان لهما أداءً جيدًا وعلى قدم المساواة وأنها أفضل من مقدر (\hat{K}_{HB}) ، ووجد أيضًا أن مقدر (\hat{K}_{KGM}) كان الأفضل من بين كل المقدرات المقترحة مما يعطي نتائج أفضل في حل مشكلة الإزدواج الخطي .
- (7) قامت دراسة (Khalaf and Shukur , 2005) [21] بتعديل معلمة (Firinguetti, 1999) [14] المعممة وقدمت بدلًا منها مقدر جديد وهو (\hat{K}_{KS}) ، الذي أظهرت نتائج الدراسة أفضليته وخاصة عندما يكون تباين حد الخطأ العشوائى كبيرًا وذلك يؤدي لتحسين مقدرات معلمات نموذج الإنحدار الخطي .
- (8) إقترحت دراسة (Habshah and Marina , 2007) [15] مقدر ريدج MM من خلال مزج مقدر MM النشط مع إنحدار ريدج لينتج مقدر Ridge MM الذى ساعد على حل مشكلتي الإزدواج الخطي والقيم الشاذة معا .
- (9) قامت دراسة (جواد & كمال ، 2009) [3] بالمقارنة بين عدة طرق لتقدير معالم الإنحدار الخطي وذلك عند وجود مشاكل الإزدواج الخطي والأخطاء التي لا تتبع التوزيع الطبيعي ومن هذه الطرق طريقة المربعات الصغرى OLS وإنحدار ريدج ومقدر الإنحرافات المطلقة الصغرى (LAD) Least Absolute Deviations ، ونموذج ريدج الموزون Weighted Ridge ، ومقدر MM النشط ، ومقدر إنحدار

ريدج المرجح بمقدر MM النشط وهو Ridge MM ، وتوصلت الدراسة إلى أن طريقة Ridge MM هي أفضل طريقة من طرق التقدير الأخرى .

(10) إعمدت دراسة (شاكر ، 2009) [6] على تقدير نموذج الإنحدار الخطى المتعدد باستخدام عدة طرق منها طريقة OLS وطريقة المربعات الصغرى المرجحة Weighted Least Squares (WLS) ، وطريقة مقدر M النشط ، وأسلوب الدمج بين نموذج WLS ونموذج M النشط وهو نموذج إنحدار Robust M Weighted وذلك لمعالجة أثر القيم الشاذة في بيانات نموذج الإنحدار ثم تمت المقارنة بينهم جميعا باستخدام معيار متوسط مربعات الخطأ Mean Squares Error ، وتوصلت الدراسة إلى أن أسلوب الدمج بين نموذج WLS ونموذج M النشط هو أكفأ نموذج في تقدير معاملات نموذج الإنحدار الخطى وبدقة عالية في ظل وجود القيم الشاذة في كل من قيم المتغيرات المستقلة والتابعة .

(11) هدفت دراسة (داوود ، 2010) [5] إلى معرفة أثر وجود بعض القيم الشاذة على نتائج تحليل الإنحدار الخطى وقد تم الكشف عن وجود القيم الشاذة بطريقة الرسم الصندوقى Box Plot وتم علاجها بواسطة طريقة الوسط الحسابي المبتور Trimmed mean ، وتوصلت الدراسة إلى عدة نتائج أهمها إن وجود القيم الشاذة في بعض المتغيرات أثر سلبا على نتائج تحليل الإنحدار حيث أنها قللت من قيمة F المحسوبة ، وكذلك قيمة معامل التحديد R^2 وزادت من قيمة متوسط مربعات الخطأ MSE وذلك مقارنة مع نتائج التحليل الجيدة التي ظهرت عقب التخلص من أثر القيم الشاذة الموجودة في بيانات متغيرات النموذج .

(12) قامت دراسة (حسن & رضا 2011) [4] بالبحث عن أنسب نموذج إنحدار خطى متعدد لتمثيل بيانات الأنواء الجوية بدراسة أهم العوامل المؤثرة في الرطوبة النسبية ، ولتحقيق هدف الدراسة تمت المقارنة بين طريقة OLS وطرق الإنحدار النشط في حالة وجود القيم الشاذة ومنها طريقة مربعات الوسيط الصغرى Least Median Squares(LMS) ومقدرات المربعات الصغرى المشذبة

Least Trimmed Squares (LTS) وتوصلت الدراسة إلى أن طريقة مربعات الوسيط الصغرى (LMS) هي أفضل طريقة لتمثيل بيانات حالة الطقس.

(13) إعتمدت دراسة (الشيخ ، 2011) [1] على استخدام نموذج إحدار ريدج في حل مشكلة الإزدواج الخطى دون اللجوء لحذف المتغيرات المستقلة المتسببة في حدوث هذه المشكلة وطبقت الدراسة على دالة الإنفاق الإستهلاكى ، وتوصلت الدراسة إلى أن مقدرات نموذج إحدار ريدج **Ridge Regression** أفضل من مقدر إحدار المربعات الصغرى العادية **OLS** .

(14) هدفت دراسة (طه ، 2014) [7] إلى معالجة مشكلة الإزدواج الخطى باستخدام نموذج إحدار ريدج وطبقت الدراسة على شركة النيل الأزرق للطباعة والتغليف من عام 1986 الى 2010 ، وتوصلت النتائج إلى أن قيمة ثابت التحيز $k=0.0900$ هي القيمة المثلى التى من شأنها أن تزيل مشكلة الإزدواج الخطى والتي بلغ عامل تضخم التباين عندها 1.322، وقد تبين أيضا أن قيم الخطأ المعيارى فى إحدار ريدج أقل من قيم الخطأ المعيارى تبعا لطريقة **OLS**.

(15) إقترحت دراسة (Pati et al., 2014 & 2016) [25 & 26] دمج مقدر ريدج مع مقدرى الإحدار النشط **LTS**, **WLS**, لإنتاج مقدرات ريدج النشطة ، كما إقترحت دراسة (Adegoke et al, 2016) [10] بدمج مقدر ريدج مع بعض المقدرات النشطة وهي مقدرات **M** , **MM** , **LTS** , **S** وذلك لتشكيل مقدرات ريدج النشطة وتوصلت الدراسة إلى حل مشكلتى الإزدواج الخطى والقيم الشاذة ورفع القدرة التفسيرية لنموذج الإحدار وخفض معامل تضخم التباين لمعاملات الإحدار المقدر.

(16) إقترحت دراسة (قاروصة ، 2019) [8] ثلاث مقدرات جديدة من مقدرات إحدار ريدج النشط وذلك لعلاج مشكلتى الإزدواج الخطى وظهور القيم الشاذة معا وهي مزيج من مقدر جديد مقترح من مقدرات ريدج مع ثلاثة مقدرات نشطة وهي المربعات الصغرى المشدبة (**LTS**) ، ومقدر **MM** ومقدر الإنحرافات المطلقة الصغرى

(LAD)، وبينت النتائج أن المقدرات الناتجة عن الدمج بين مقدرات ريدج ومقدرات الإنحدار النشط تعطي أفضل النتائج في حل مشكلتي الأزواج الخطي والمشاهدات الشاذة معا عن النماذج الأخرى .

أما عن الدراسة الحالية تختلف عن الدراسات السابقة في إنها ستقوم بعقد مقارنة بين ست مقدرات لنموذج الإنحدار الخطي المتعدد وهي OLS, Robust M, Robust MM, HKB Ridge, Robust Ridge M, Robust Ridge MM وإختيار أفضلها من حيث القدرة التفسيرية و التنبؤية لنموذج الإنحدار الخطي ، هذا بالإضافة بأنه لم توجد دراسة من الدراسات السابقة طبقت هذه المقدرات معا وقامت بإختيار أفضلها للتنبؤ بإنتاج وإستهلاك المياه النقية في محطات المياه التابعة للشركة القابضة للمياه في مصر .

رابعا : منهجية الدراسة وتحليل البيانات

(1/4) نموذج إنحدار المربعات الصغرى العادية Ordinary Least Squares يعد نموذج الإنحدار الخطي المتعدد هو أحد النماذج الإحصائية اللازمة لدراسة العلاقة بين المتغير التابع (Y) وعدد من المتغيرات المستقلة (X) والتي يأخذ نموذجها الشكل التالي:-

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i \quad i = 1, 2, \dots, n \quad (1)$$

حيث يعبر Y_i عن المتغير التابع ، $X'S$ تمثل المتغيرات المستقلة، $(\beta_0, \beta_1, \beta_2, \dots, \beta_k)$ تعبر عن معاملات الإنحدار، ε_i تعبر عن الخطأ العشوائي للمشاهدة رقم i ، $i=1,2,\dots,n$ ، وحيث أن n تمثل عدد المشاهدات ويمكن حساب مقدر $\hat{\beta}$ بطريقة OLS بإستخدام المصفوفات من المعادلة التالية:

$$\hat{\beta}_{LS} = (X'X)^{-1} X'Y \quad (2)$$

حيث أن $(X'X)$ تمثل مصفوفة معاملات الارتباط البسيطة بين كل زوج من المتغيرات المستقلة ، ويمثل المتجه $(X'Y)$ معاملات الارتباط بين المتغير التابع وكل متغير مستقل وكما يمكن حساب تباين مقدر المربعات الصغرى من المعادلة التالية:-

$$MSE(\hat{\beta}_{LS}) = Cov(\hat{\beta}_{LS}) = \sigma^2 (X'X)^{-1} \quad (3)$$

وفى الواقع العملى نلاحظ أن هناك مشاكل عديدة فى نموذج الإنحدار الخطى ومنها مشكلة الإزدواج الخطى **Multicollinearity** بين المتغيرات المستقلة ومشكلة المشاهدات الشاذة **Outliers** التى توجد فى بيانات المتغيرات التابعة والمستقلة ، وتؤدى مشكلة الإزدواج الخطى إلى تضخم فى تباينات مقدرات معاملات الإنحدار (Habshah and Marina, 2007) [15] وبالتالي لى يتم تخفيض هذه التباينات نستخدم نموذج إنحدار ريدج والذى يحسن من دقة تقديرات معاملات الإنحدار باستخدام طريقة OLS (Samkar & Aplu , 2010) [33] .

نموذج انحدار ريدج (2/4) Ridge Regression Model

يعتبر مقدر إنحدار ريدج (Hoerl & Kennard, 1970 a & b) [16 & 17] هو الأكثر شيوعا وإستخداما فى التعامل مع مشكلة الإزدواج الخطى وهو يهدف إلى تقليل حجم تباين تقديرات المربعات الصغرى عن طريق إدخال كمية ضئيلة من التحيز ، ويأخذ المقدر الصيغة العامة التالية :-

$$\hat{\beta}_{GR} = (X'X + KI)^{-1} X'Y \quad (4)$$

حيث $K \geq 0$ تسمى معلمة ريدج Ridge أو معلمة التحيز Bais Parameter وعندما تكون k مساوية للصفر فإن مقدر ريدج يصبح هو نفسه مقدر OLS ويأخذ متوسط مربعات خطأ مقدر ريدج $MSE(\hat{\beta}_R)$ المعادلة التالية :-

$$MSE(\hat{\beta}_R) = \sigma^2 \sum_{i=1}^p \frac{\lambda_i}{(\lambda_i + k)^2} + k^2 \beta' (X'X + kI_p)^{-2} \beta \quad (5)$$

حيث λ_i هي القيم المميزة Eigen Values للمصفوفة $(X'X)$ وعلى الرغم من أن مقدر ريدج هو مقدر متحيز ، إلا أنه عند قيمة معينة K يكون له أقل قيمة MSE مقارنة بمقدر OLS ومع ذلك فإن $MSE(\hat{\beta}_R)$ يعتمد على المعالم (σ^2, β, k) والتي لا يمكن حسابها عمليا وبالتالي يجب تقدير المعلمة (K) من البيانات المشاهدة والتي تعتمد على أخذ الوسط التوافقي Harmonic Mean لمعلمة ريدج المعممة (Horel et. al. ,1975) [18] وتأخذ معادلته الشكل التالي:-

$$\hat{K}_{HKB} = \frac{ps^2}{\sum_{i=1}^p \hat{\alpha}_i^2} = \frac{ps^2}{\sum_{i=1}^p \hat{\beta}_i^2} = \frac{ps^2}{\hat{\beta}'\hat{\beta}} \quad (6)$$

حيث تمثل $\hat{\alpha}, \hat{\beta}$ مقدرات OLS لمتجهات المعالم α, β ويصبح كذلك s^2 هو مقدر OLS لتباين حد الخطأ العشوائي والذي يأخذ الشكل التالي مع المعالم $\hat{\alpha}, \hat{\beta}$:-

$$s^2 = (Y - X\hat{\alpha})'(Y - X\hat{\alpha}) / (n - p) \quad (7)$$

$$s^2 = (Y - X\hat{\beta})'(Y - X\hat{\beta}) / (n - p) \quad (8)$$

أما عن المشكلة الثانية التي يمكن أن تواجه نموذج الإنحدار الخطي هي وجود المشاهدات الشاذة في المتغيرات التابعة والمستقلة وهي المشاهدات التي تحرف عن علاقة الإنحدار الموقفة بغالبية البيانات (Rousseuw & Leroy, 1987) [32] وهي تبدو غير متسقة مع بقية المشاهدات ، وغالبا ماتوجد المشاهدات الشاذة في البيانات لأسباب كثيرة منها أخطاء في القياس أو تجميع البيانات من مجتمعات مختلفة (Barnett & Lewis, 1994) [12] وهي تؤدي لوجود تقديرات متحيزة مع تضخم في تباينات مقدرات معاملات الإنحدار أو المتغير التابع المقدر ، أو يمتد أثرها لتضخم كليهما معا (Chatterjee & Handi, 2006) [13] وللتغلب على مشكلة المشاهدات الشاذة يتم استخدام نموذج الإنحدار لأن مقدراته تتميز بأنها أقل تأثرا بالقيم الشاذة من مقدر OLS .

Robust Regression Model (3 / 4) نموذج الإحدار النشط

قد أشار (Ronchetti, 1987) [29] إلى أن الهدف من استخدام نموذج الإحدار النشط هو إختيار النموذج الذي يوفق غالبية البيانات مع الأخذ في الإعتبار أن الأخطاء قد لاتتبع التوزيع الطبيعي المعتدل ، وقد تم إقتراح العديد من مقدرات الإحدار النشط ومنها مقدر (LTS) ومقدر Robust M ومقدر Robust MM ومقدرا الانحرافات المطلقة الصغرى (LAD)، ومقدر المربعات الوسيطة الصغرى (LMS)، ومقدر Robust S (Rousseeuw, 1984 & 1985) [30 & 31] وستقتصر الدراسة الحالية على تطبيق مقدرى Robust M و Robust MM على نموذج الإحدار الخطى لأنهما من أفضل مقدرات نموذج الإحدار النشط ويتم توضيح طريقة حسابهما كمايلى :-

- قدم (Huber ,1972a) [20] مقدر Robust M Reg. ويتم به حساب تقديرات النموذج بواسطة تخفيض مجموع حاصل ضرب الأخطاء العشوائية مرجحا بدالة الأوزان Huber Function ويحسب المقدر كمايلى :-

$$\hat{\beta}_M = \min_B \sum_{i=1}^n \rho(r_i) \quad (9)$$

كما قدم (Yohai , 1987) [34] مقدر Robust MM ، وهو أكثر إنتشارا لأنه أكثر المقدرات مقاومة للمشاهدات الشاذة فى إتجاه كل من متغير التابع والمتغيرات المستقلة . ويرجع الاسم MM إلى إستخدام مقدر M فى أكثر من مرحلة لحساب التقديرات النهائية ، ويتم الحصول على مقدر MM على عدة مراحل وتنتهى إلى الحصول على النهاية الصغرى المطلقة لمايلى :-

$$L(\hat{\beta}_{MM}) = \sum_{i=1}^n \rho_1 \left(\frac{r_i}{\hat{\sigma}_{(0)}} \right) x_i = 0 \quad (10)$$

وعندما توجد مشكلتي الإزدواج الخطى والقيم الشاذة معا في نموذج الإنحدار يكون من الأفضل إستخدام نموذج الدمج إنحدار ريدج النشط **Robust Ridge Reg.** كما يتضح في النقاط التالية .

(4/4) نموذج إنحدار ريدج النشط **Robust Ridge Regression Model**

إقتراح (Pafaffenberger & Dielman, 1985) [28] بدمج مقدر إنحدار ريدج مع مقدر M النشط لينتج مقدر **Robust Ridge M Reg.** كما يلي:

$$\hat{\beta}_{RRM} = (X^t X + K_M I)^{-1} X^t Y \quad (11)$$

وتأخذ \hat{K}_M الصيغة التالية :-

$$\hat{K}_M = \frac{pS_M^2}{\hat{\beta}_M^t \hat{\beta}_M} \quad (12)$$

حيث أن $\hat{\beta}_M$ هو مقدر M النشط لمتجه الإنحدار $(\hat{\beta}_M, S_M^2)$ وحيث أن :-

$$S_M^2 = \frac{(Y - X\hat{\beta}_M)^t (Y - X\hat{\beta}_M)}{(n-p)} \quad (13)$$

كما إقتراح (Habshah & Marina, 2007) [15] بدمج مقدر ريدج مع مقدر **MM** النشط لينتج مقدر **Robust Ridge MM Reg.** ويأخذ الصيغة التالية :-

$$\hat{\beta}_{RRMM} = (X^t X + K_{MM} I)^{-1} X^t Y \quad (14)$$

وتأخذ \hat{K}_{MM} الصيغة التالية :-

$$\hat{K}_{MM} = \frac{pS_{MM}^2}{\hat{\beta}_{MM}^t \hat{\beta}_{MM}} \quad (15)$$

حيث أن $\hat{\beta}_{MM}$ هو مقدر **MM** النشط لمتجه الإنحدار $(\hat{\beta}_{MM}, S_{MM}^2)$ وحيث أن :-

$$S_{MM}^2 = \frac{(Y - X\hat{\beta}_{MM})^t (Y - X\hat{\beta}_{MM})}{(n-p)} \quad (16)$$

ومن الجدير بالذكر أن \hat{K}_M, \hat{K}_{MM} هما نفسهما K_{HKB} كما يتضح في المعادلتين (11 ، 14) مع وجود إختلافيين هما إستخدام مقدرى **MM** ، **M** في تقدير متجه

المعالم β بدلا من مقدر OLS ثم يتم تعديل S^2 باستخدام تقديرات MM ، M بدلا من استخدام تقديرات OLS كما يتضح في المعادلتين (13 ، 16)، وتهدف هذه التغييرات إلى تخفيض تأثير المشاهدات الشاذة على القيمة المختارة لمعلمة ريدج .
(4 / 5) الدراسة التطبيقية للنماذج الإحصائية

يتضمن هذا الجزء تطبيق للنماذج الإحصائية حيث نبدأ بتقدير وتقييم نموذج الإتحاد الخطي المتعدد بطريقة OLS ثم إجراء الإختبارات التشخيصية للنموذج ثم نقوم بالكشف عن مدى وجود مشكلتي الإزدواج الخطي والقيم الشاذة في بيانات دوال الإنتاج والإستهلاك لكميات المياه النقية في محطات المياه التابعة للشركة القابضة ولتحليل البيانات سيتم استخدام برنامجي (Eviews10, NCSS7.1).
تتمثل دالة إستهلاك المياه النقية في المتغير التابع وهو إجمالي كمية المياه المستهلكة (WC_{it}) بالمتر المكعب والمتغيرات المستقلة التي تتمثل في كميات المياه المستهلكة في الوحدات السكنية (HU_{it})، والمحال التجارية (CS_{it}) ، والورش والمصانع (WF_{it})، وشركات السياحة (TC_{it})، وأماكن أخرى (OP_{it}) بالإضافة إلى فاقد الشبكات (NL_{it}) وتأخذ دالة إستهلاك المياه الشكل التالي :-

$$WC_{it} = HU_{it} + CS_{it} + WF_{it} + TC_{it} + OP_{it} + NL_{it} + \varepsilon_{it} \quad (17)$$

وتتمثل دالة إنتاج المياه في المتغير التابع وهو إجمالي كمية المياه المنتجة (WP_{it}) بالمتر المكعب والمتغيرات المستقلة تتمثل في العوامل المؤثرة على كمية إنتاج المياه في الشركة القابضة وهي عدد ساعات التشغيل بالمحطة يوميا (OH_{it}) ، وعدد أيام التشغيل بالمحطة سنويا (OD_{it})، القدرة الفعلية للمحطة (AC_{it}) ، القدرة التصميمية للمحطة (DC_{it}) ، عدد المحطات (NS_{it}) ، وتأخذ دالة إنتاج المياه الشكل التالي :-

$$WP_{it} = OH_{it} + OD + AC_{it} + DC_{it} + NS_{it} + \varepsilon_{it} \quad (18)$$

(1/5/4) تقديرات نموذج الإحدار لدالة إستهلاك المياه النقية .

يوضح الجدول رقم (3) نتائج تقديرات نموذج المربعات الصغرى العادية كمايلي : -

| Variable | Coeff. | Std. Error | T - stat | Prob. | VIF |
|-------------------------|---------|--------------------------|----------|--------|-------|
| C | 1516613 | 461316 | 3.288 | 0.006 | - |
| CS | 12.392 | 2.856 | 4.339 | 0.001 | 6.119 |
| HU | 1.025 | 0.761 | 1.348 | 0.201 | 2.941 |
| NL | -0.870 | 0.388 | -2.241 | 0.043 | 6.889 |
| OP | 2.173 | 0.766 | 2.835 | 0.014 | 2.278 |
| TC | 0.814 | 1.209 | 0.673 | 0.513 | 3.086 |
| WF | 3.564 | 1.168 | 3.051 | 0.009 | 3.590 |
| R-squared | 0.949 | Mean dep.var. | | 6E+06 | |
| Adjusted R- Squ. | 0.9165 | SD. dep.var. | | 2E+06 | |
| S.E. of reg. | 566592 | Akaike inf. crit. | | 29.602 | |
| Sum squ. resid. | 4E+12 | Schwarz crit. | | 29.95 | |
| Log likelihood | -289.02 | Hannan-Qui.cr. | | 29.67 | |
| F-statistic | 35.775 | Durb.-Wat.stat | | 2.0666 | |
| Prob(Fstatistic) | 0.000 | | | | |

- يوضح الجدول السابق أنه توجد علاقة معنوية بين إجمالي كمية المياه المستهلكة (WC_{it}) بالمتري المكعب و كل من كميات المياه المستهلكة في المحال التجارية (CS_{it}) و الورش والمصانع (WF_{it}) ، وفاقد الشبكات (NL_{it}) . والإماكن الأخرى (OP_{it})،
- توجد علاقة معنوية للنموذج ككل حيث بلغت قيمة (F) المحسوبة 35.78 وقيمة P-Value أقل من 5 % وبلغت قيمة معامل التحديد (0.943)، وهو مايعكس إرتفاع القدرة التفسيرية للنموذج وهذا يعنى أن المتغيرات المستقلة التى يتضمنها النموذج تفسر نسبة 94.3 % من التغيرات التى تحدث فى المتغير التابع وأن 5.7 % من التغيرات يرجع لعوامل أخرى منها الخطأ العشوائى .

- يتضح من الجدول رقم (3) أيضا خلو النموذج المقدر من الارتباط الذاتي حيث بلغت قيمة إحصاء **Durbin –Watson** حوالي 2.07 ونلاحظ من الجدول السابق أيضا أن قيم معامل تضخم التباين **VIF** لجميع المتغيرات المستقلة أقل من 10 وبالتالي لا يوجد ازدواج خطى بين المتغيرات المستقلة .
- يوضح الجدول رقم (4) الإختبارات التشخيصية للنموذج حيث يوضح إختبار **Breusch – Godfrey LM Test** أنه لا يوجد مشكلة الارتباط الذاتي بين البواقي كما أوضحه إختبار ديرين واتسن (DW) حيث كانت قيمة **Obs*R-squared** تساوي 1.078 وأن قيمة **Prob.Chi-Sq.(2)** تساوي 0.5834 وهي أكبر من 5% مما يعنى قبول فرض العدم الذي يؤكد عدم وجود ارتباط ذاتي بين البواقي .

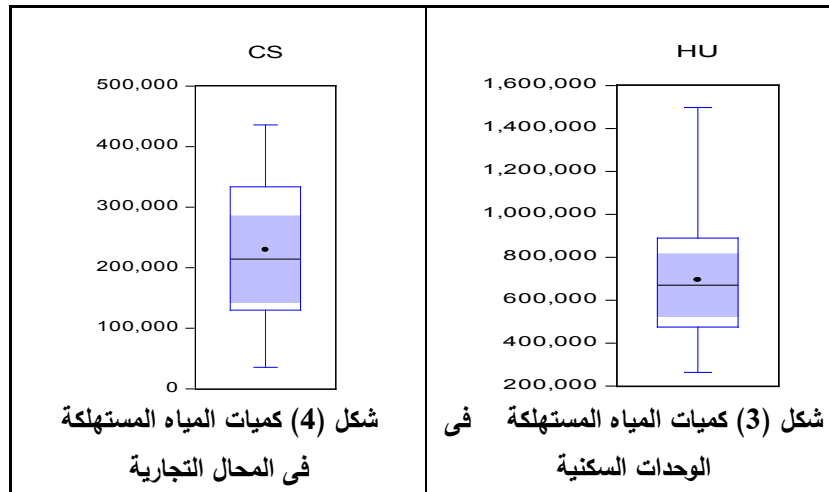
جدول (4) إختبارات الارتباط الذاتي وعدم ثبات التباين والطبيعية للبواقي

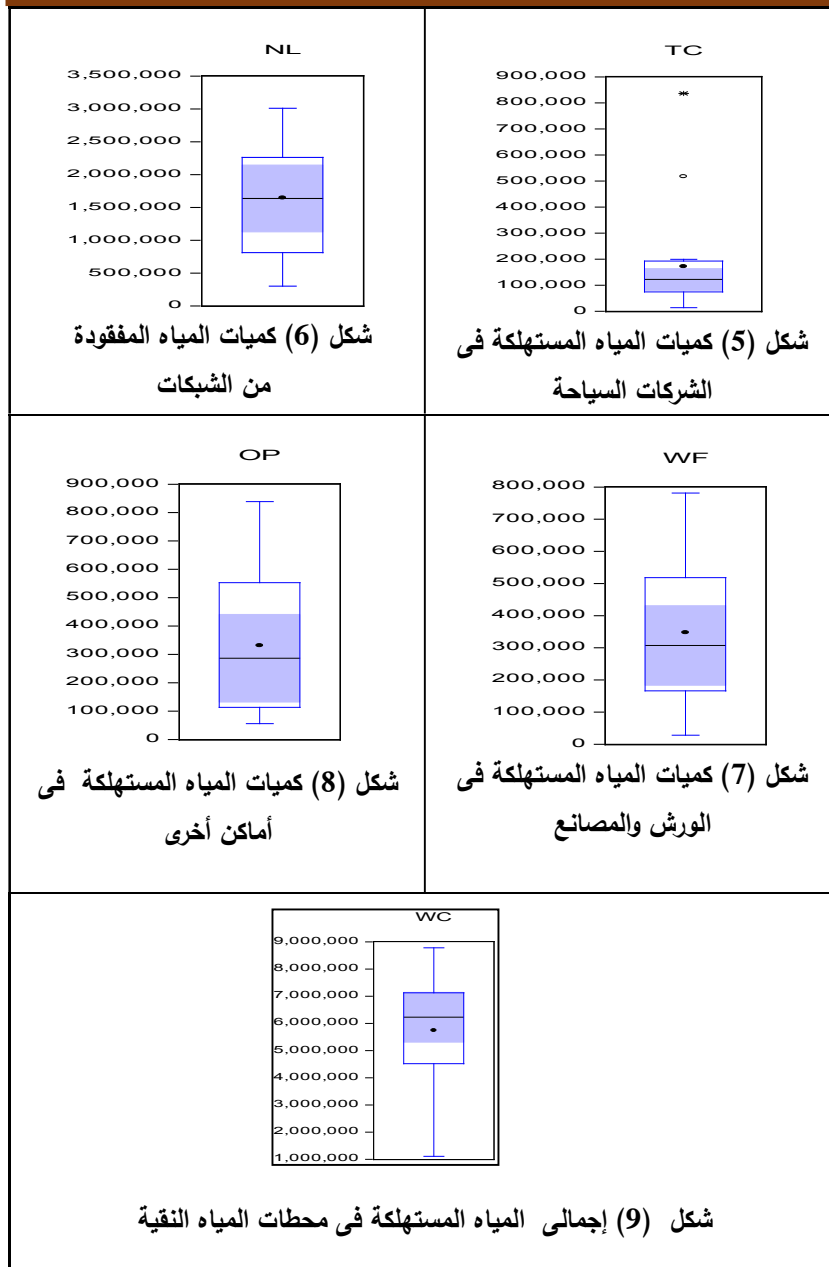
| Breusch-Godfrey Serial Correlation LM Test : | | | |
|--|----------|----------------------|---------|
| F-statistic | 0.313235 | Prob. F(2,11) | 0.7374 |
| Obs*R-squared | 1.077661 | Prob. Chi-Square(2) | 0.5834 |
| Heteroskedasticity Test: Breusch-Pagan-Godfrey | | | |
| F-statistic | 1.145352 | Prob. F(6,13) | 0.3907 |
| Obs*R-squared | 6.916337 | Prob. Chi-Square(19) | 0.3287 |
| Scaled expla.SS | 3.758819 | Prob. Chi-Square(19) | 0.7093 |
| Normality Test | | | |
| Jarque – Bera | 1.143605 | Prob.obs(20) | 0.56451 |

- كما يبين جدول(4) نتائج إختبار **Breusch –Pagan –Godfrey** بأن النموذج المقدر لا يواجه مشكلة عدم ثبات التباين بين البواقي حيث كانت قيمة **Obs*R-squared** تساوي 6.916 وأن قيمة **Prob. Chi-Sq.(2)** تساوي 0.3287 وهي أكبر من 5% مما يعنى قبول فرض العدم الذي يؤكد عدم وجود مشكلة عدم ثبات التباين للبواقي ، وكذلك يبين الجدول السابق نتائج إختبار " Jarque – Bera " والذي يوضح أن سلسلة البواقي للنموذج المقدر

لا تتبع التوزيع الطبيعي حيث بلغت قيمة إختبار $JB = 1.144$ وكانت قيمة P-Value تساوى 0.56451 وهى أعلى من 5 % .

- أما عن وجود القيم الشاذة وكيفية إكتشافها فمن خلال الأشكال البيانية التالية وباستخدام الرسم الصندوقى Box Plot فقد تبين وجود قيمتين شاذتين فى بيانات شركات السياحة (TC_{it}) كما يظهر فى الشكل رقم (3) الذى سيؤدى لوجود تقديرات متحيزة و تضخم فى تباينات المقدرات معاملات الإنحدار الأمر الذى يلزم معه إستخدام مقدرات نماذج الإنحدار النشط وإنحدار ريدج وذلك لإزالة أثر القيم الشاذة على البيانات وتحسين جودة توفيق النموذج والحصول على مقدرات غير متحيزة ، أما عن المتغيرات الباقية المتمثلة فى إجمالى كمية المياه المستهلكة (WC_{it}) وكمية المياه المستهلكة فى كل من الوحدات السكنية (HU_{it}) والمحال التجارية (CS_{it})، والورش والمصانع (WF_{it}) ، والأماكن الأخرى (OP_{it}) بالإضافة إلى فاقد الشبكات (NL_{it}) فلا يوجد بها أى قيم شاذة كما توضحه باقى الأشكال التالية :-





يوضح الجدول رقم (5) مقارنة لمقدرات دالة إستهلاك المياه وذلك بإستخدام طريقة المربعات الصغرى العادية OLS ومقدرات نموذج الإنحدار النشط ونموذج إنحدار ريدج وأسلوب الدمج بين نموذجي الإنحدار النشط وإنحدار ريدج كمايلي: -
جدول (5) مقارنة مقدرات دالة إستهلاك المياه النقية بإستخدام نماذج

(OLS,Robust M ,Robust MM, Ridge,Rob.Ridge M, Rob.Ridge MM)

| Variable | Const. | CS | HU | NL | OP | TC | WF |
|-----------------------------|------------|-----------------|------------------|--------------|-----------------------|------------------------|--------|
| OLS | | | | | | | |
| Coff. | 1516613 | 12.39 | 1.025 | -0.87 | 2.173 | 0.814 | 3.564 |
| P.value | 0.006 | 0.001 | 0.201 | 0.043 | 0.014 | 0.513 | 0.009 |
| Robust M Reg. | | | | | | | |
| Coff. | 2008030 | 13.20 | 0.509 | -0.976 | 2.1589 | 0.7654 | 3.3796 |
| P.value | 0.000 | 0.000 | 0.314 | 0.000 | 0.000 | 0.342 | 0.000 |
| Robust MM Reg. | | | | | | | |
| Coff. | 2409684 | 13.33 | 0.31 | 2.295 | 2.295 | 1.071 | 3.512 |
| P.value | 0.000 | 0.000 | 0.471 | 0.000 | 0.000 | 0.117 | 0.000 |
| Ridge | | | | | | | |
| Coff. | 1522037 | 12.07 | 1.024 | -0.812 | 2.1233 | 60.72 | 3.576 |
| P.value | 0.000 | 0.000 | 0.200 | 0.000 | 0.000 | 0.510 | 0.000 |
| Robust Ridge M Reg. | | | | | | | |
| Coff. | 2008438 | 12.83 | 0.519 | -0.915 | 2.105 | 0.6773 | 3.4078 |
| P.value | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Robust Ridge MM Reg. | | | | | | | |
| Coff. | 16689 | 16.31 | 2.943 | -1.205 | 1.184 | 0.639 | 4.275 |
| P.value | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.200 | 0.000 |
| | OLS | Robust M | Robust MM | Ridge | Robust Ridge M | Robust Ridge MM | |
| F-test | 35.78 | 447.6 | 579.8 | 33.03 | 391.6 | 366.7 | |
| Prob. (F-Stat) | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | |
| R² | 90.94 | 0.856 | 0.769 | 0.9384 | 0.995 | 0.994 | |
| RMSE | 456801 | 486142 | 563495 | 588249 | 164116 | 214996 | |

يوضح الجدول السابق مقارنة بين العديد من مقدرات نماذج الإنحدار ، وقد تبين أن أفضل نموذج معنوى له أعلى قدرة تفسيرية وتنبؤية هو نموذج Robust Ridge MM Regression وتبين منه مايلى :-

- أن هناك علاقة معنوية بين إجمالى كمية المياه المستهلكة فى محطات المياه النقية (WC_{it}) وبين كل من كميات المياه المستهلكة فى الوحدات السكنية (HU_{it}) ، والمحال التجار (CS_{it}) ، والورش والمصانع (WF_{it}) ، وشركات السياحة (TC_{it}) ، والأماكن الأخرى (OP_{it}) بالإضافة لفاقد الشبكات (NL_{it}) كمايلى :-

$$WC_{it} = 2008438 + 12.83CS_{it} + 0.519HU_{it} - 0.915NL_{it} + 2.105OP_{it} + 0.677TC_{it} + 3.408WF_{it} \quad (19)$$

- توجد علاقة معنوية للنموذج ككل حيث بلغت قيمة (F) المحسوبة 391.6 وقيمة P-Value أقل من 5 % وبلغت قيمة معامل التحديد (0.995) ، وهو يعكس ارتفاع القدرة التفسيرية للنموذج ويعنى أن المتغيرات المستقلة التى يتضمنها النموذج تفسر نسبة 99.5 % من التغيرات التى تحدث فى المتغير التابع وأن 0.5 % من التغيرات يرجع لعوامل أخرى منها الخطأ العشوائى ، كما تبين من النموذج أيضا أن له أعلى قدرة تنبؤية ويتضح ذلك من معيار RMSE والنسبة بلغت قيمته 164116 ، وبالتالي يصلح هذا النموذج فى للتنبؤ بكميات المياه المستهلكة فى الشركة القابضة للمياه من عام 2020 حتى عام 2030

(2/5/4) تقديرات نموذج الإنحدار لدالة إنتاج المياه النقية .

يوضح الجدول رقم (6) نتائج تقديرات نموذج المربعات الصغرى العادية OLS لدالة إنتاج المياه النقية والذى يبين أنه توجد علاقة معنوية بين إجمالى كمية المياه المنتجة فى محطات المياه (WP_{it}) بالمتر المكعب و عدد محطات المياه النقية (NS_{it}) التابعة للشركة القابضة للمياه .

جدول (6) نتائج تقديرات نموذج OLS لدالة إنتاج المياه النقية

| Variable | Coeff. | Std. Err. | T-Stat. | Prob. | VIF |
|---------------|----------|------------------------|----------|----------|--------|
| C | -1.2E+07 | 4708210 | -2.47394 | 0.0268 | - |
| AC | 41.56691 | 20.82817 | 1.995706 | 0.0658 | 2.0633 |
| DC | 5.491984 | 2.747156 | 1.999153 | 0.0654 | 1.1475 |
| NS | 2202.807 | 399.0095 | 5.520688 | 0.0001 | 6.0030 |
| OD | 27659.95 | 20675.79 | 1.337794 | 0.2023 | 3.5722 |
| OH | 241821.1 | 284797.5 | 0.849098 | 0.4101 | 12.880 |
| R-Squared | 0.919881 | Mean Depend. Var. | | 6797547 | |
| Adj. R Squ. | 0.891267 | S.D. Dependent var | | 2445222 | |
| S.E.of Regr. | 806305.2 | Akaike Info. Criterion | | 30.28164 | |
| Sum Squ. Res. | 9.10E+12 | Schwarz Criterion | | 30.58036 | |
| Log Likelih. | -296.816 | Hannan-Quinn Criter. | | 30.33995 | |
| F-Statistic | 32.14792 | Durbin-Watson stat | | 2.024806 | |
| Prob(F-stat) | 0.000 | | | | |

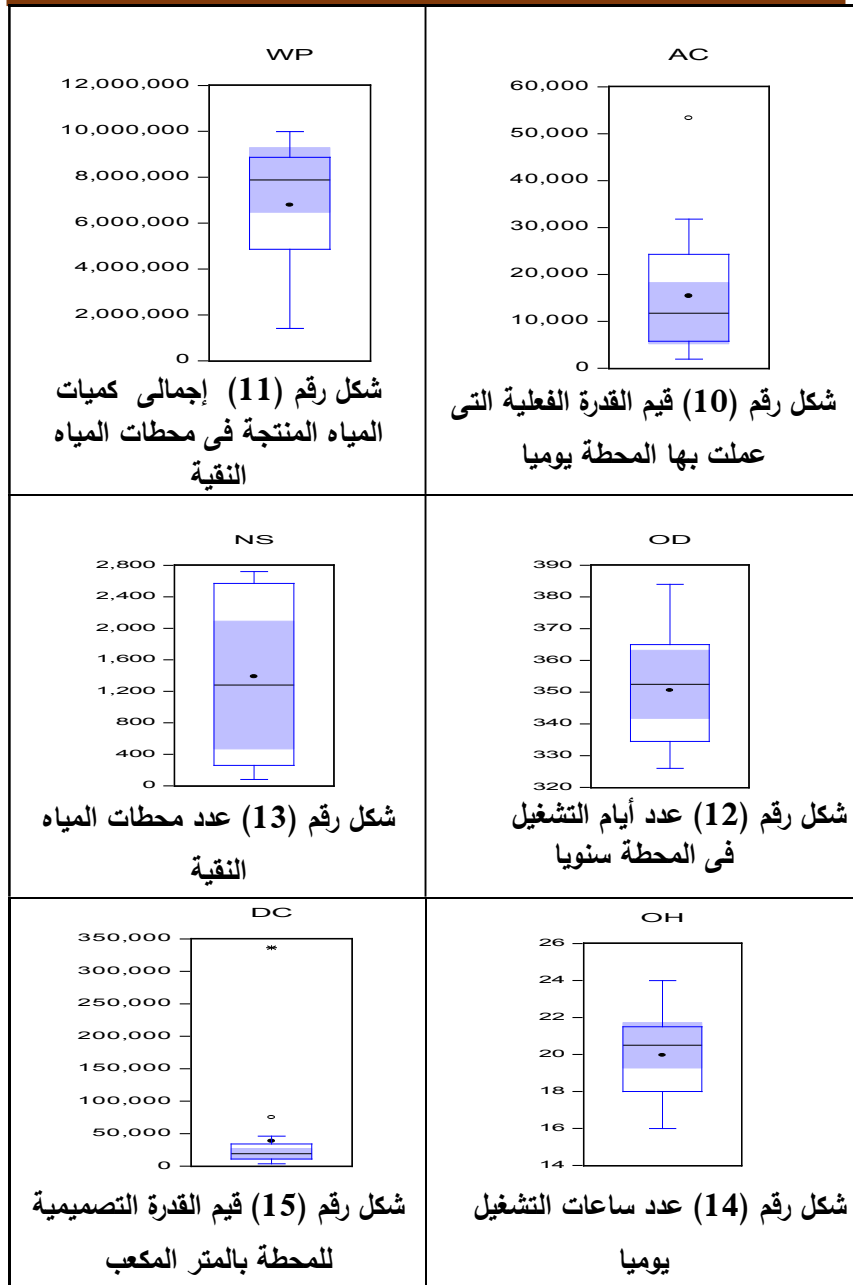
- يوضح الجدول السابق معنوية نموذج الإنحدار ككل حيث بلغت قيمة (F) المحسوبة 32.15 وقيمة P-Value أقل من 5 % ، كما بلغت قيمة معامل التحديد (0.919) وهو يعكس إرتفاع القدرة التفسيرية للنموذج ويعنى أن المتغيرات المستقلة تفسر نسبة 92 % تقريبا من التغيرات التي تحدث في المتغير التابع وأن 8 % من التغيرات ترجع لعوامل أخرى منها الخطأ العشوائى.
- يتضح من الجدول رقم (6) أيضا خلو النموذج المقدر من الارتباط الذاتى للبواقي حيث بلغت قيمة إحصاء Durbin –Watson حوالى 2.03 ، كما نلاحظ من الجدول السابق أن قيم معامل تضخم التباين VIF لجميع المتغيرات المستقلة أقل من 10 فيما عدا متغير عدد ساعات التشغيل للمحطة (OH_{it}) حيث بلغ معامل تضخم التباين 12.88 وهو أعلى من 10 مما يعنى وجود مشكلة الإزدواج الخطى بين المتغيرات المستقلة .

- يوضح جدول (7) الإختبارات التشخيصية الذي تبين فيها عدم وجود مشكلة الإرتباط الذاتي بين البواقي ويتضح ذلك من نتيجة إختبار **Breusch Godfrey LM Test** حيث كانت قيمة **Obs*R-squared** تساوى 2.081 بقيمة محسوبة (2). **Prob.Chi-Sq.** تساوى 0.3533 مما يعنى قبول فرض عدم الذي يؤكد عدم وجود إرتباط ذاتى بين البواقي .

جدول (7) إختبارات الإرتباط الذاتي وعدم ثبات التباين والطبيعية للبواقي

| Breusch-Godfrey Serial Correlation LM Test : | | | |
|--|----------|---------------------|--------|
| F-Statistic | 0.696798 | Prob. F(2,12) | 0.5173 |
| Obs*R-squared | 2.080988 | Prob.Chi-Square(2) | 0.3533 |
| Heteroskedasticity Test: Breusch-Pagan-Godfrey | | | |
| F-Statistic | 0.319541 | Prob. F(5,14) | 0.893 |
| Obs*R-Squared | 2.04864 | Prob. Chi-Square(5) | 0.8424 |
| Scaled Explained SS | 1.895811 | Prob.Chi-Square(5) | 0.8634 |
| Normality Test | 10.02074 | Prob.Obs.(20) | 0.0067 |
| Jarque – Bera | | | |

كما يبين جدول (7) نتائج إختبار **Breusch-Pagan-Godfrey** بأن النموذج المقدر لا يواجه مشكلة عدم ثبات التباين بين البواقي **Heteroskedasticity** حيث كانت قيمة **Obs*R-squared** تساوى 2.049 بقيمة محسوبة (5). **Prob.Chi-Sq.** تساوى 0.8634 وكذلك يبين الجدول السابق نتائج إختبار " Jarque – Bera " والذي يوضح أن سلسلة البواقي للنموذج المقدر لا تتبع التوزيع الطبيعي حيث بلغت قيمة إختبار $JB = 10.02$ بقيمة محسوبة **P- Value** تساوى 0.0067 ، ويرجع ذلك لوجود القيم الشاذة فى البيانات فى كل من متغير القدرة الفعلية للمحطة (AC_{it}) ، ومتغير القدرة التصميمية (DC_{it}) والتي يجب إزالتها بإستخدام نموذج الإنحدار النشط ، أما متغير إجمالى كمية المياه النقية المنتجة (WP_{it}) وباقى المتغيرات المستقلة مثل عدد ساعات التشغيل يوميا فى المحطة (OH_{it}) ، عدد أيام التشغيل (OD_{it}) ، عدد محطات المياه النقية (NS_{it}) تخلو من القيم الشاذة ويتضح ذلك من الأشكال البيانية التالية :-



يتضح من النتائج السابقة والأشكال البيانية وجود مشكلتي الإزدواج الخطى والقيم الشاذة في بعض البيانات الأمر الذي يؤدي لوجود تضخم في تباينات مقدرات معالم النموذج الأمر الذي إستخدام مقدرات نماذج الإنحدار النشط وريدج ونماذج الدمج بينهما لمعالجة هاتين المشكلتين معا .

جدول (8) مقارنة مقدرات دالة إنتاج المياه النقية بإستخدام نماذج (OLS, Robust M ,Robust MM, Ridge,Rob.Ridge M, Rob.Ridge MM)

| Variable | Const. | AC | DC | NS | OD | OH |
|-----------------------------|------------|-----------------|------------------|--------------|---------------------|---------------------|
| OLS | | | | | | |
| Coff. | -1 E+07 | 41.5669 | 5.492 | 2202.81 | 27660 | 241821 |
| P.value | 0.027 | 0.066 | 0.065 | 0.000 | 0.202 | 0.410 |
| Robust M Reg. | | | | | | |
| Coff. | -9510398 | 45.0524 | 4.92861 | 2247.011 | 15504 | 350112 |
| P.value | 0.009 | 0.005 | 0.019 | 0.000 | 0.328 | 0.109 |
| Robust MM Reg. | | | | | | |
| Coff. | -5998677 | 40.0287 | 4.73483 | 2102.5 | 8104 | 319007 |
| P.value | 0.080 | 0.008 | 0.018 | 0.000 | 0.590 | 0.123 |
| Ridge | | | | | | |
| Coff. | -1.1E+07 | 40.4873 | 5.5139 | 2128.99 | 29594 | 195054 |
| P.value | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.410 |
| Robust Ridge M Reg. | | | | | | |
| Coff. | -9142074 | 43.614 | 4.96 | 2163 | 71804 | 293936 |
| P.value | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Robust Ridge MM Reg. | | | | | | |
| Coff. | -5668570 | 38.679 | 4.765 | 2024 | 10554 | 265967 |
| P.value | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | OLS | Robust M | Robust MM | Ridge | Rob. Ridge M | Rob. Rid. MM |
| F-test | 32.148 | 257.50 | 259.43 | 29.78 | 391.14 | 410.9 |
| Prob. (F-Stat) | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| R ² | 0.920 | 0.859 | 0.776 | .914 | .9929 | 0.9932 |
| RMSE | 674603 | 698213 | 736683 | 835129 | 223423 | 207474 |

يوضح الجدول (8) مقارنة مقدرات دالة إنتاج المياه النقية باستخدام نماذج (OLS, Robust M ,Robust MM,Ridge, Rob.Ridge M, Rob. Ridge MM) وقد تبين من الجدول أن أفضل نموذج معنوي له أعلى قدرة تفسيرية وتنبؤية هو نموذج **Robust Ridge MM Regression** وتبين منه أن هناك علاقة معنوية بين إجمالي كمية المياه المنتجة بالمتري المكعب (WP_{it}) و عدد ساعات التشغيل يوميا (OH_{it}) ، وعدد أيام التشغيل سنويا (OD_{it}) ، والقدرة الفعلية للمحطة (AC_{it}) ، والقدرة التصميمية للمحطة (DC_{it}) ، وعدد المحطات (NS_{it}) ، وتأخذ دالة إنتاج المياه الشكل التالي : -

$$WP_{it} = -5668570 + 38.679AC_{it} + 4.765DC_{it} + 2024NS_{it} + 10554OD_{it} + 265967OH_{it} \quad (20)$$

كما يوجد علاقة معنوية للنموذج ككل حيث بلغت قيمة (F) المحسوبة 410.9 وأن مستوى المعنوية أقل من 5 % كما بلغت قيمة معامل التحديد (0.993) مما يعني أن المتغيرات المستقلة تفسر نسبة 99.3 % من التغيرات التي تحدث في المتغير التابع وأن 0.7 % من التغيرات يرجع لعوامل أخرى منها الخطأ العشوائي ، ووجد أيضا أن قيمة ثابت التحيز التي تزيل مشكلة الإزدواج الخطي $K = 0.005$ ، كما تبين أن النموذج له أعلى قدرة تنبؤية وتبين ذلك من معيار $RMSE$ والتي بلغت قيمته 207474 ، وبالتالي يصلح هذا النموذج للتنبؤ بكميات المياه المنتجة في الشركة القابضة لمياه الشرب من عام 2020 حتى عام 2030 .

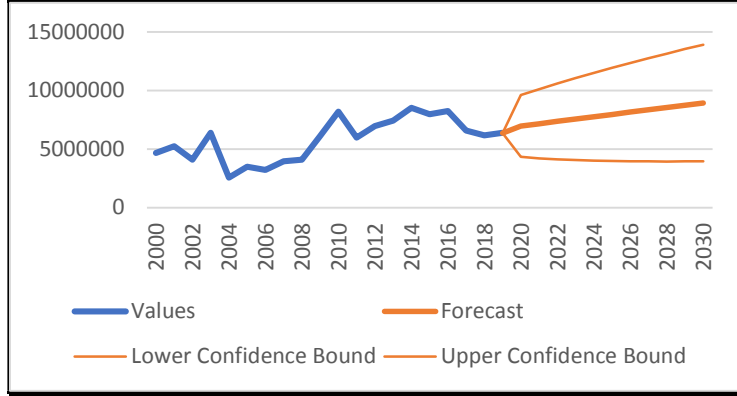
(6 / 4) التنبؤ

يوضح الجدول رقم (9) القيم التنبؤية وفترات الثقة لكميات المياه المستهلكة والمنتجة في محطات المياه النقية التابعة للشركة القابضة للمياه في مصر وذلك في الفترة من 2020 إلى 2030 على النحو التالي :-

جدول (9) القيم التنبؤية لإجمالي كميات المياه المستهلكة والمنتجة في محطات المياه النقية التابعة للشركة القابضة للمياه في مصر

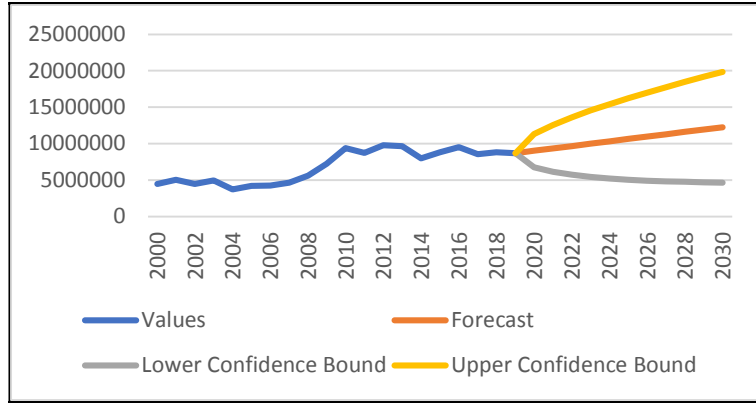
| Year | Water Consumed (WC _{it}) | | | Water Produced (WP _{it}) | | |
|------|------------------------------------|---------|-----------|------------------------------------|---------|----------|
| | Forecast | LCL | UCL | Forecast | LCL | UCL |
| 2020 | 6963429 | 4321013 | 9605846 | 9025881 | 6745126 | 11306635 |
| 2021 | 7161419 | 4205925 | 10116914 | 9349427 | 6125565 | 12573288 |
| 2022 | 7359409 | 4119883 | 10598936 | 9672973 | 5723906 | 13622039 |
| 2023 | 7557399 | 4055809 | 11058989 | 9996519 | 5435009 | 14558029 |
| 2024 | 7755389 | 4009089 | 11501689 | 10320064 | 5218100 | 15422029 |
| 2025 | 7953379 | 3976516 | 119302423 | 10643610 | 5052265 | 16234956 |
| 2026 | 8151369 | 3955755 | 12346984 | 10967156 | 4925080 | 17009232 |
| 2027 | 8349359 | 3945044 | 12753674 | 11290702 | 4828452 | 17752952 |
| 2028 | 8547349 | 3943016 | 13151683 | 11614248 | 4756762 | 18471734 |
| 2029 | 8745339 | 3948582 | 13542097 | 11937794 | 4705917 | 19169671 |
| 2030 | 8943329 | 3960860 | 13925799 | 12261340 | 4672832 | 19849847 |

يتضح من الجدول السابق أنه بإستخدام القيم المقدرة من نموذج Robust Ridge M Regression في الفترة من عام 2000 حتى عام 2019 يمكننا التنبؤ بإجمالي كميات المياه المستهلكة في محطات المياه النقية من عام 2020 حتى عام 2030 بإستخدام السلاسل الزمنية والتي يوضحها الجدول رقم (9) وحيث يتضح من الجدول أنه من المتوقع زيادة كميات المياه المستهلكة من 6963429 متر مكعب عام 2020 حتى تصل إلى 8943329 متر مكعب عام 2030 ويمكن تمثيل هذه القيم التنبؤية في الشكل البياني التالي :-



شكل (16) القيم التنبؤية لكميات المياه المستهلكة في محطات المياه النقية

يتضح من الجدول (9) أيضا أنه باستخدام القيم المقدرة من نموذج Robust Ridge MM Regression في الفترة من عام 2000 حتى عام 2019 يمكننا التنبؤ بإجمالي كميات المياه المنتجة من عام 2020 حتى عام 2030 والتي يوضحها الجدول رقم (9) حيث يبين أنه من المتوقع زيادة كميات المياه المنتجة من 9025881 متر مكعب عام 2020 حتى تصل إلى 12261340 متر مكعب عام 2030 ، ويمكن تمثيل هذه القيم التنبؤية كما في الشكل البياني التالي :-



شكل (17) القيم التنبؤية لكميات المياه المنتجة في محطات المياه النقية

خامسا : النتائج والتوصيات

قامت الدراسة الحالية بدراسة العوامل المؤثرة على كميات المياه المستهلكة والمنتجة في محطات المياه النقية التابعة للشركة القابضة للمياه في الفترة من عام 2000 إلى 2019 وذلك بهدف الوصول إلى أفضل نموذج للتنبؤ بإنتاج وإستهلاك المياه في هذه الشركة في الفترة من 2020 إلى عام 2030 وقد تمت المقارنة بين كل من نموذج المربعات الصغرى العادية (OLS) ، ومقدرات نموذج الإنحدار النشط Robust Regression ، ومقدرات نموذج إنحدار ريدج Robust Ridge Regression بالإضافة لأسلوب الدمج بين النموذجين Robust Ridge Regression وذلك من حيث القدرة التفسيرية والتنبؤية ، وقد قامت الدراسة على إختبار أربع فروض إحصائية وكانت النتائج كمايلي :-

- (1) رفض الفرض الأول لأن نموذج Robust Ridge M Regression قد بين أن هناك علاقة ذات دلالة إحصائية معنوية بين إجمالي كمية إستهلاك المياه في الشركة القابضة (WC_{it}) وبين كمية المياه المستهلكة في الوحدات السكنية (HU_{it})، والمحال التجارية (CS_{it}) ، والورش والمصانع (WF_{it}) ، وشركات السياحة (TC_{it})، وأماكن أخرى (OP_{it}) هذا بالإضافة لفاقد الشبكات (NL_{it})
- (2) رفض الفرض الثاني لأن نموذج Robust Ridge MM Regression قد بين أن هناك علاقة ذات دلالة إحصائية معنوية بين إجمالي كمية المياه المنتجة بالمتري المكعب (WP_{it}) وبين عدد ساعات التشغيل يوميا (OH_{it}) ، وعدد أيام التشغيل في المحطة (OD_{it})، والقدرة الفعلية للمحطة (AC_{it}) ، و القدرة التصميمية لها (DC_{it}) بالإضافة إلى عدد المحطات (NS_{it}) .

(3) قبول الفرض الثالث وهو أن نموذج الدمج **Robust Ridge Regression** هو أفضل نموذج للتنبؤ بكميات إنتاج وإستهلاك المياه النقية وذلك لإرتفاع قدرته التفسيرية والتنبؤية عن نموذج المربعات الصغرى العادية (OLS)، ونموذج الإنحدار **Robust Regression**، ونموذج إنحدار ريدج **Ridge Regression**

(4) رفض الفرض الرابع الذى يتوقع حدوث زيادة فى كميات إستهلاك المياه فى الشركة القابضة للمياه فى مصر وإنخفاض كميات المياه المنتجة خلال فترة التنبؤ وذلك لأن نتائج الدراسة قد بينت أنه من المتوقع أن تكون هناك زيادة فى كميات المياه المستهلكة والمنتجة معا فى محطات المياه النقية فى الشركة القابضة .

(5) طبقا لنتائج الدراسة يوصى الباحث بإستخدام نموذج إنحدار ريدج **Robust Ridge Regression** فى التنبؤ بكميات المياه المستهلكة والمنتجة فى الشركة القابضة فى المستقبل لإرتفاع قدرته التنبؤية عن النماذج الأخرى .

(6) نظرا لما أظهرته النتائج أيضا من حدوث زيادة فى إستهلاك المياه فى الشركة القابضة فى الفترة المستقبلية فمن الضرورى التحول من ثقافة وفرة المياه التى كانت سائدة فى مصر إلى ثقافة ندرة المياه التى أصبحت واقعا ، ولذلك يجب العمل على وجود مزيد من حملات التوعية الإعلامية بهدف تقليل من العوامل المسببة لها والتى أظهرت الدراسة معنويتها ألا وهى كميات المياه المستهلكة فى كل من الوحدات السكنية (HU_{it}) والمحال التجارية (CS_{it}) والورش والمصانع (WF_{it}) ، وشركات السياحة (TC_{it})، والأماكن الأخرى (OP_{it}) وفاقد الشبكات (NL_{it}) وذلك لمدى أهميتها بالنسبة للمجتمع ككل .

(7) يجب على الشركة القابضة والشركات التابعة لها القيام بتجديد شبكات المياه القديمة وصيانتها بصورة دورية لتقليل فاقد الشبكات (NL_{it}) والإستفادة منه فى المستقبل .

(8) بينت نتائج الدراسة زيادة كميات المياه المنتجة فى محطات المياه فى المستقبل ومن ثم يجب الحفاظ على العوامل المسببة لذلك والتي أظهرت النتائج معنويتها ألا وهى عدد ساعات التشغيل يوميا (OH_{it}) ، وعدد أيام التشغيل سنويا (OD_{it}) ، والقدرة الفعلية التى تعمل بها المحطة (AC_{it}) ، هذا بالإضافة للقدرة التصميمية لها (DC_{it}) ، وزيادة عدد محطات المياه (NS_{it}).

(9) ضرورة تضافر جهود كافة المعنيين بقضايا المياه فى الدولة ورفع الوعى بتحديات قطاع المياه وكيفية مواجهتها على كافة المستويات السياسية والتنفيذية والشعبية وأن ذلك أصبح ضرورة ملحة وليس خيارا حتى يمكن تحسين كفاءة إستخدام الموارد المائية المحدودة فى مصر والحفاظ عليها وذلك خصوصا لمواجهة قضية سد النهضة فى أثيوبيا التى قد تؤثر سلبا على إحتياجات مصر فى المستقبل .

أولاً : المراجع العربية :-

- (1) الشيخ ، ساوس (2014). معالجة مشكلة الإزدواج الخطى بإستخدام إنحدار الحرف : دراسة تطبيقية على دالة الإنفاق الإستهلاكى فى الجزائر خلال الفترة 1970 - 2011 ، مجلة الحقيقة ، جامعة أدرار ، الجزائر، العدد 29، ص 34 - 55 .
- (2) المجلس العربى للمياه (2015) . ورشة عمل خبراء التدريب حول دليل إعداد الإستراتيجيات المائية وخطط العمل مع تضمينها تأثير التغيرات المناخية ، ص 1-14 .
- (3) جواد ، نزار مصطفى و كمال ، غفران إسماعيل (2009) . مقارنة طرق تقدير معالم نموذج الإنحدار فى حالة ظهور مشكلة التعدد الخطى والقيم الشاذة ، مجلة العلوم الإدارية والإقتصادية ، العدد (55) ، العراق ، بغداد ، ص 153 - 166 .
- (4) حسن ، تارا أحمد ورضا ، مهدى صابر (2011). إستخدام الإنحدار الحصين لإيجاد أنسب نموذج لتمثيل بيانات الأنواء الجوى فى مدينة أربيل خلال الفترة 1998 - 2020 ، مجلة العلوم الإقتصادية والإدارية ، العدد (89) ، الجامعة المستنصرية ، ص 195 - 213 .
- (5) داوود ، ألاء عبد الستار (2010) ، تأثير القيم الشاذة ونقطة الأصل على نتائج تحليل الإنحدار ، مجلة تكريت للعلوم الصرفة ، المجلد 15 ، العدد (1)، ص 310 - 319 .

(6) شاكر ، صالح مؤيد (2009) . تحسين أسلوب M الحصين في تقدير معلمات نموذج الإنحدار الخطى المتعدد ، المجلة العراقية للعلوم الإحصائية ، العدد (16) ، ص 219-242 .

(7) طه ، إمتثال ميرغنى محمد أحمد (2014) . معالجة مشكلة التداخل الخطى المتعدد لشركة النيل الأزرق للتغليف والطباعة باستخدام إنحدار الحرف (1986 - 2010) ، رسالة ماجستير، كلية العلوم ، جامعة السودان للعلوم والتكنولوجيا .

(8) قاروصة ، أحمد (2019) . دراسة مقارنة لبعض مقدرات إنحدار ريدج المتينة ، مجلة كلية التجارة للبحوث العلمية ، كلية التجارة ، جامعة الإسكندرية ، العدد الأول ، المجلد السادس والخمسون ، يناير ، ص 1-27 .

(9) وزارة الموارد المائية والرى (2017) . مشروع دعم الخطة القومية للموارد المائية ممول من برنامج الإتحاد الأوروبى لإصلاح قطاع المياه فى مصر ، المرحلة الثانية .

ثانيا : المراجع الأجنبية :-

(10) Adegoke , A.S., Adewuyi , E., Ayinde,K., and Lukman,A.F. (2016) , **A comparative study of some robust ridge and Liu estimators** , Science World Journal , 11, 4 , 16-20 .

(11) Askin, R.G., and Montgomery,D.C. (1980), **Augmented robust estimators** , Technometrics , 22, PP 333-341.

- (12) Barnett, V., and Lewis, T. (1994), **Outliers in statistical data**, 3rd ed., Wiley : Chicester, UK .
- (13) Chatterjee, S., and Handi, A.S. (2006), **Regression Analysis by Example** , Wiley : New Jersey.
- (14) Firinguetti, L. (1999), **Generalized ridge regression estimator and its finite Sample properties** , Communications in Statistics: Theory and Methods, 28(5) , 1217 – 1229.
- (15) Habshah, M., and Marina, z. (2007), **A Simulation study on ridge regression estimators in Presence of outliers and multicollinearity** , Journal Teknologi, Universiti Teknologi Malaysia ,47, 59-74.
- (16) Hoerl, A.E., and Kennard, R.W. (1970 a), **Ridge regression : Biased estimation for non-orthogonal Problems** , Technometrics, 12, 55-67.
- (17) Hoerl, A.E., and Kennard, R.W. (1970 b), **Ridge regression : Application to non-orthogonal Problems** , Technometrics, 12, 69-82.

- (18) Hoerl , A.E., and Kennard , R.W.(1975) , **Ridge regression: some simulations** ,Communications in Statistics ,4, 105-123 .
- (19) Holland, P.W.(1973), **Weighted ridge and robust methods**, BERN Working Paper Series , Working Paper No. 11, PP1-19.
- (20) Huber, P.J.(1972a) , **Robust statistics** , Ann. Math. Statist., 43, 1041 – 1067 .
- (21) Khalaf, G. and Shukur ,G.(2005), **Choosing ridge parameter for regression problem** , Communications in Statistics : Theory and Methods, 34, 1177 – 1182.
- (22) Kibira ,B.M.G.(2003), **Performance of some new ridge regression estimators**, Communications in Statistics : Simulation and Computation, 32, 419 – 435.
- (23) Lawless, J.F., and Wang, P.(1976), **A Simulation study of ridge and other regression estimators**, Communications in Statistics : Theory and Methods , 14, 1589 - 1604 .

- (24) Nomura, M.(1988), **On the almost unbiased ridge regression estimation**, Communications in Statistics : Simulation and computation , 17(3) , 729 - 743 .
- (25) Pati, K.D.,Adanan,R.,Rasheed ,B.A.,and Alias, M.J.(2014), **Ridge least trimmed squares estimators in the presence of multicollinearity and outliers** , Nature Science, 12,12,1-8 .
- (26) Pati, K .D., Adanan,R.,Rasheed ,B.A.,and Alias, M.J. (2016), **Estimation paramerters using bisquare weighted robust ridge regression (BRLTS) estimators in the presence of multicollinearity and outlie** , AIP Conference Proceedings 1750,060-028;doi:10.1631/1.4954 633.
- (27) Pfaffenberger, R.C. and Dielman,T.E. (1984), **A modified ridge regression estimators using the least absolute value criterion in the multiple linear regression model**, Proceedings of the American Institute for Decision Sciences, Toronto , 791-793.
- (28) Pfaffenberger,R.C. and Dielman,T.e.(1985). **A Comparison of robust ridge estimators** , Business Economics Sectios Proceedings of the American Statistical Association , 631-635.

- (29) Ronchetti , E.m. (1987). **Robust $C(\alpha)$ -type tests for linear models** , Sankhya. A,49 , 1-16 .
- (30) Rousseeuw,P.J.(1984), **Least median of squares regression**, Journal of the American Statistical Association, 79, 871- 880.
- (31) Rousseeuw , P. J. (1985) , **Multivariate estimation with high breakdown point in Mathematical Statistics and Applications** , B (ed . by W. W . Wertz) , Dordrecht, Netherlands : Reidel Publishing Company , 283-297 .
- (32) Rousseeuw,P. J., and Leroy ,A.M. (1987) , **Robust regression and outlier detection** ,Wiley : New York .
- (33) Samkar , H., and Alpu, O.(2010) , **Ridge regression based on some robust estimators** , Journal of Modern Applied Statistical Methods ,9,2, Article 17.
- (34) Yohai, V. J. (1987), **High breakdown point and high efficiency robust estimates for regression** , Ann. Statistics ., 15, 642 – 656 .

