





معاونت پژوهش و فن آوری  
به نام خدا  
منشور اخلاق پژوهش

بایاری از خداوند سبحان و اعتماد به این که عالم محضر خداست و همواره ناظر بر اعمال انسان و به منظور پاس داشت مقام بلند دانش و پژوهش و نظریه  
اهمیت جایگاه دانشگاه در اعتلای فرهنگ و تمدن بشری، ماداشجویان و اعضاء هیئت علمی واحدهای دانشگاه آزاد اسلامی متعهد می گردیم اصول زیر را  
در انجام فعالیت های پژوهشی مد نظر قرار داده و از آن تخطی نکنیم:

- ۱- اصل حقیقت جویی: تلاش در راستای پی جویی حقیقت و وفاداری به آن و دوری از هرگونه پنهان سازی حقیقت.
- ۲- اصل رعایت حقوق: التزام به رعایت کامل حقوق پژوهشگران و پژوهشیدگان (انسان، حیوان و نبات) و سایر صاحبان حق.
- ۳- اصل مالکیت مادی و معنوی: تعهد به رعایت کامل حقوق مادی و معنوی دانشگاه و کلیه بهکاران پژوهش.
- ۴- اصل منافع ملی: تعهد به رعایت مصالح ملی و در نظر داشتن پیشبرد و توسعه کشور در کلیه مراحل پژوهش.
- ۵- اصل رعایت انصاف و امانت: تعهد به اجتناب از هرگونه جانب داری غیر علمی و حفاظت از اموال، تجهیزات و منابع در اختیار.
- ۶- اصل رازداری: تعهد به صیانت از اسرار و اطلاعات محرمانه افراد، سازمان ها و کشور و کلیه افراد و نهاد های مرتبط با تحقیق.
- ۷- اصل احترام: تعهد به رعایت حریم ها و حرمت ها در انجام تحقیقات و رعایت جانب تقد و خودداری از هرگونه حرمت شکنی.
- ۸- اصل ترویج: تعهد به رواج دانش و اشاعه نتایج تحقیقات و انتقال آن به بهکاران علمی و دانشجویان به غیر از مواردی که منع قانونی دارد.
- ۹- اصل برائت: التزام به برائت جویی از هرگونه رفتار غیر حرفه ای و اعلام موضع نسبت به کسانی که حوزه علم و پژوهش را به شائبه های غیر علمی می

آلایند.



معاونت پژوهش و فن آوری  
به نام خدا  
تعهد اصالت رساله یا پایان نامه تحصیلی

اینجانب داود اورنگ دانش آموخته مقطع کارشناسی ارشد در رشته آبیاری زه کشی که در تاریخ ۹۳/۳/۷ از پایان نامه خود تحت عنوان " مقایسه روش های باز سازی داده های مفقود مطالعه موردی رودخانه کر " با کسب نمره دفاع نموده ام بدینوسیله متعهد می شوم:

این پایان نامه حاصل تحقیق و پژوهش انجام شده توسط اینجانب بوده و در مواردی که از دستاوردهای علمی و پژوهشی دیگران (اعم از پایان نامه، کتاب، مقاله و ...) استفاده نموده ام، مطابق ضوابط و رویه موجود، نام منبع مورد استفاده و سایر مشخصات آنرا در فهرست مربوطه ذکر و درج کرده ام.

این پایان نامه قبلاً برای هیچ مدرک تحصیلی (هم سطح، پایین تر یا بالاتر) در سایر دانشگاه ها و موسسات آموزش عالی ارائه نشده است.

چنانچه بعد از فراغت از تحصیل، قصد استفاده و هرگونه بهره برداری اعم از چاپ کتاب، ثبت اختراع و ... از پایان نامه داشته باشم، از حوزه معاونت پژوهشی واحد مجوزهای مربوطه را اخذ نمایم.

چنانچه در هر مقطع زمانی خلاف موارد فوق ثابت شود، عواقب ناشی از آن را می پذیرم و دانشگاه آزاد اسلامی واحد مرودشت مجاز است با اینجانب مطابق ضوابط و مقررات رفتار نموده و در صورت ابطال مدرک تحصیلی ام هیچگونه ادعایی نخواهم داشت.

نام و نام خانوادگی:

تاریخ و امضاء



دانشگاه آزاد اسلامی  
واحد مرودشت  
دانشکده کشاورزی - گروه مهندسی آب

پایان نامه برای دریافت درجه کارشناسی ارشد (M.Sc.)  
گرایش: آبیاری و زهکشی

عنوان:

**مقایسه روش های باز سازی داده های مفقود مطالعه موردی رودخانه کر**

استاد راهنما:

دکتر هما رزمخواه

استاد مشاور:

دکتر محمد هادی فتاحی

نگارش:

داود اورنگ

بهار ۱۳۹۳



## صور تجلسه دفاع از پایان نامه کارشناسی ارشد (M.Sc.)

نام و نام خانوادگی دانشجو: **داود اورنگ** در تاریخ: **۱۳۹۳/۳/۷** رشته: **آبیاری و زهکشی**  
از پایان نامه خود با عنوان: **مقایسه روش های باز سازی داده های مفقود مطالعه موردی رودخانه**  
**کر**

با درجه و نمره دفاع نموده است.

نام و نام خانوادگی اعضای هیات داوری سمت امضاء اعضای هیات داوری

۱- دکتر هما رزمخواه استاد راهنما

۲- دکتر محمد هادی فتاحی استاد مشاور

۳- دکتر امین رستمی راوری استاد داور

مراتب فوق مورد تایید است.  
مدیر/معاون پژوهشی  
مهر و امضاء

## تقدیر و تشکر

سرکار خانم دکتر هما رزمخواه و جناب آقای دکتر محمد هادی فتاحی و جناب آقای دکتر امین رستمی شما روشنایی بخش تاریکی جان هستی و ظلمت اندیشه را نور می بخشی. چگونه سپاس گویم مهربانی و لطف تو را که سرشار از عشق و یقین است. چگونه سپاس گویم تأثیر علم آموزی تو را که چراغ روشن هدایت را بر کلبه ی محقر وجودم فروزان ساخته است. آری در مقابل این همه عظمت و شکوه تو مرا نه توان سپاس است و نه کلام وصف.

## تقدیم

و بعد از مدتها، پس از پیمودن راههای فراوان که با حضور شیرین اساتید عزیزم، با راهنماییها و دغدههای فراوانشان نگاههای پدر مادرم، با چشمهای پر از برق شوق، و زیبایی حضور خواهرم در کنارم، که خستگیهای این راه را به امید و روشنی راه تبدیل کرده و امیدوارم بتوانم در آینده ی نزدیک جوابگوی این همه محبت آنها باشم ...

اکنون، با احترام فراوان برای این همه تلاش این عزیزان برای موفقیت من ....  
این پایان نامه را به پدر و مادرم، اساتید عزیز و خواهر مهربانم تقدیم میکنم.  
امیدوارم قادر به درک زیباییهای وجودشان باشم.

## فهرست مطالب

عنوان	شماره صفحه
<b>فصل اول: کلیات تحقیق</b>	
۱-۱ مقدمه.....	۳
۲-۱ فرضیه.....	۵
۳-۱ اهداف تحقیق.....	۵
<b>فصل دوم: ادبیات تحقیق و پیشینه تحقیق</b>	
۱-۲ مروری بر تحقیقات انجام شده.....	۸
۲-۲ منطقه مورد مطالعه.....	۱۲
۱-۲-۲ مقدمه.....	۱۲
۲-۲-۲ موقعیت جغرافیایی.....	۱۳
۳-۲-۲ فیزیو گرافی منطقه.....	۱۶
۴-۲-۲ هواشناسی.....	۱۶
۵-۲-۲ بادها و توده های هوا.....	۱۷
۶-۲-۲ رژیم آبدهی.....	۱۸
<b>فصل سوم: روش اجرای تحقیق</b>	
۱-۳ مواد و روش ها.....	۲۰
۱-۳-۱ مقدمه.....	۲۰
۲-۱-۳ آزمون نرمال بودن در SPSS.....	۲۰
۱-۲-۱-۳ آزمون کولموگروف-اسمیرنوف.....	۲۰
۱-۲-۳ آزمون همگنی داده ها.....	۲۰
۲-۲-۳ نمودار پراکنندگی.....	۲۱
۳-۲-۳ بررسی چولگی داده ها.....	۲۱
۴-۲-۳ معرفی روش های بازسازی.....	۲۲
۱-۴-۲-۳ رگرسیون به روش معمولی.....	۲۲
۲-۴-۲-۳ خط پایدار کندال - تیل.....	۲۳
۳-۴-۲-۳ خط همبستگی طبیعی.....	۲۴



۲۵	..... خط همبستگی بنیانی
	<b>فصل چهارم: تجزیه و تحلیل داده ها</b>
۲۸	..... ۱-۴ مقدمه
۲۸	..... ۲-۴ بررسی نرمال بودن داده ها
۳۳	..... ۳-۴ بررسی همگنی داده ها
۳۷	..... ۴-۴ نمودار Box plot
۳۸	..... ۵-۴ بررسی چولگی داده ها
۴۳	..... ۶-۴ ارزیابی روش های بازسازی داده های مفقود
۴۳	..... ۱-۶-۴ RMSE
۴۳	..... ۲-۶-۴ MAE
۴۴	..... ۳-۶-۴ MBE
۴۴	..... ۷-۴ روش OLS
۵۲	..... ۸-۴ روش KTRL
۵۷	..... ۹-۴ روش LOC
۶۲	..... ۱۰-۴ روش RLOC
۶۶	..... ۱۱-۴ بررسی نتایج آزمون های ارزیابی
	<b>فصل پنجم: نتیجه گیری و پیشنهادات</b>
۶۸	..... بحث و نتایج
۶۹	..... منابع و مآخذ

## فهرست جداول

شماره صفحه	عنوان
۱۷.....	جدول ۱-۲ فاکتورهای فیزیکی حوزه آبریز رودخانه کر در بالادست سد درودزن.....
۱۹.....	جدول ۲-۲ مشخصات ایستگاه های آبنجی حوزه رودخانه کر.....
۲۸.....	جدول ۱-۴ آزمون کولموگروف-اسمیرنوف ایستگاه چمریز CL ۲۶.....
۲۹.....	جدول ۲-۴ آزمون کولموگروف-اسمیرنوف ایستگاه دهکده سفید CL.....
۲۹.....	جدول ۳-۴ آزمون کولموگروف-اسمیرنوف ایستگاه درودزن CL.....
۳۰.....	جدول ۴-۴ آزمون کولموگروف-اسمیرنوف ایستگاه حسن آباد CL.....
۳۰.....	جدول ۵-۴ آزمون کولموگروف-اسمیرنوف ایستگاه پل خان CL.....
۳۱.....	جدول ۶-۴ آزمون کولموگروف-اسمیرنوف ایستگاه چمریز دبی.....
۳۱.....	جدول ۷-۴ آزمون کولموگروف-اسمیرنوف ایستگاه درود زن دبی.....
۳۲.....	جدول ۸-۴ آزمون کولموگروف-اسمیرنوف ایستگاه پل خان دبی.....
۳۲.....	جدول ۹-۴ آزمون کولموگروف-اسمیرنوف ایستگاه دهکده سفید دبی.....
۳۳.....	جدول ۱۰-۴ آزمون کولموگروف-اسمیرنوف ایستگاه حسن آباد دبی.....
۳۳.....	جدول ۱۱-۴ ران تست ایستگاه پل خان CL.....
۳۴.....	جدول ۱۲-۴ ران تست ایستگاه حسن آباد CL.....
۳۴.....	جدول ۱۳-۴ ران تست ایستگاه درود زن CL.....
۳۵.....	جدول ۱۴-۴ ران تست ایستگاه چمریز CL.....
۳۵.....	جدول ۱۵-۴ ران تست ایستگاه دهکده سفید CL.....
۳۵.....	جدول ۱۶-۴ ران تست ایستگاه چمریز دبی.....
۳۶.....	جدول ۱۷-۴ ران تست ایستگاه دهکده سفید دبی.....
۳۶.....	جدول ۱۸-۴ ران تست ایستگاه درود زن دبی.....
۳۶.....	جدول ۱۹-۴ ران تست ایستگاه حسن آباد دبی.....
۳۷.....	جدول ۲۰-۴ ران تست ایستگاه پل خان دبی.....
۴۴.....	جدول ۲۱-۴ معیارهای ارزیابی روش OLS در بازسازی داده های CL ، ایستگاه حسن آباد- پل خان.....

- جدول ۴-۲۲ معیارهای ارزیابی روش OLS در بازسازی داده های CL ، ایستگاه حسن آباد- درودزن..... ۴۵
- جدول ۴-۲۳ معیارهای ارزیابی روش OLS در بازسازی داده های CL ، ایستگاه چمریز-درودزن . ۴۶
- جدول ۴-۲۴ معیارهای ارزیابی روش OLS در بازسازی داده های CL ، ایستگاه چمریز-دهکده سفید..... ۴۷
- جدول ۴-۲۵ معیارهای ارزیابی روش OLS در بازسازی داده های دبی، ایستگاه دهکده سفید- چمریز ..... ۴۸
- جدول ۴-۲۶ معیارهای ارزیابی روش OLS در بازسازی داده های دبی، ایستگاه درودزن- چمریز..... ۴۹
- جدول ۴-۲۷ معیارهای ارزیابی روش OLS در بازسازی داده های دبی، ایستگاه درودزن- پلخان ..... ۵۰
- جدول ۴-۲۸ معیارهای ارزیابی روش OLS در بازسازی داده های دبی، ایستگاه حسن آباد- پلخان ..... ۵۱
- جدول ۴-۲۹ معیارهای ارزیابی روش OLS در بازسازی داده های دبی، ایستگاه حسن آباد- درودزن ..... ۵۲
- جدول ۴-۳۰ ارزیابی روش KTRL در بازسازی داده های CL ایستگاه حسن آباد-پلخان..... ۵۳
- جدول ۴-۳۱ ارزیابی روش KTRL در بازسازی داده های CL ایستگاه چمریز-درودزن ..... ۵۳
- جدول ۴-۳۲ ارزیابی روش KTRL در بازسازی داده های CL ایستگاه چمریز-درودزن ..... ۵۴
- جدول ۴-۳۳ ارزیابی روش KTRL در بازسازی داده های CL ایستگاه دهکده سفید-چمریز ..... ۵۴
- جدول ۴-۳۴ ارزیابی روش KTRL در بازسازی داده های دبی ایستگاه چمریز-دهکده سفید ..... ۵۵
- جدول ۴-۳۵ ارزیابی روش KTRL در بازسازی داده های دبی ایستگاه چمریز-درودزن ..... ۵۵
- جدول ۴-۳۶ ارزیابی روش KTRL در بازسازی داده های دبی ایستگاه پلخان-درودزن ..... ۵۶
- جدول ۴-۳۷ ارزیابی روش KTRL در بازسازی داده های دبی ایستگاه پلخان-حسن آباد ..... ۵۶
- جدول ۴-۳۸ ارزیابی روش KTRL در بازسازی داده های دبی درودزن-حسن آباد ..... ۵۷
- جدول ۴-۳۹ ارزیابی روش LOC در بازسازی داده های CL ایستگاه پلخان-حسن آباد ..... ۵۷
- جدول ۴-۴۰ ارزیابی روش LOC در بازسازی داده های CL ایستگاه درودزن-حسن آباد ..... ۵۸
- جدول ۴-۴۱ ارزیابی روش LOC در بازسازی داده های CL ایستگاه درودزن-چمریز ..... ۵۸
- جدول ۴-۴۲ ارزیابی روش LOC در بازسازی داده های CL ایستگاه دهکده سفید-چمریز ..... ۵۹
- جدول ۴-۴۳ ارزیابی روش LOC در بازسازی داده های دبی ایستگاه دهکده سفید-چمریز ..... ۵۹

- جدول ۴-۴۴ ارزیابی روش LOC در بازسازی داده های دبی ایستگاه درودزن-چمریز ..... ۶۰
- جدول ۴-۴۵ ارزیابی روش LOC در بازسازی داده های دبی ایستگاه درودزن-پلخان ..... ۶۰
- جدول ۴-۴۶ ارزیابی روش LOC در بازسازی داده های دبی ایستگاه حسن آباد-پلخان ..... ۶۱
- جدول ۴-۴۷ ارزیابی روش LOC در بازسازی داده های دبی ایستگاه حسن آباد-درودزن ..... ۶۱
- جدول ۴-۴۸ ارزیابی روش RLOC در بازسازی داده های CL (حسن آباد-پلخان) ..... ۶۲
- جدول ۴-۴۹ ارزیابی روش RLOC در بازسازی داده های CL (حسن آباد-درودزن) ..... ۶۲
- جدول ۴-۵۰ ارزیابی روش RLOC در بازسازی داده های CL (چمریز-درودزن) ..... ۶۳
- جدول ۴-۵۱ ارزیابی روش RLOC در بازسازی داده های CL (چمریز-دهکده) ..... ۶۳
- جدول ۴-۵۲ ارزیابی روش RLOC در بازسازی داده های دبی (چمریز-دهکده) ..... ۶۴
- جدول ۴-۵۳ ارزیابی روش RLOC در بازسازی داده های دبی (چمریز-درودزن) ..... ۶۴
- جدول ۴-۵۴ ارزیابی روش RLOC در بازسازی داده های دبی (پلخان-درودزن) ..... ۶۵
- جدول ۴-۵۵ ارزیابی روش RLOC در بازسازی داده های دبی (پلخان-حسن آباد) ..... ۶۵
- جدول ۴-۵۶ ارزیابی روش RLOC در بازسازی داده های دبی (درودزن-حسن آباد) ..... ۶۶

## فهرست نمودارها

عنوان	شماره صفحه
نمودار ۱-۴ box-plot پارامتر دبی	۳۷
نمودار ۲-۴ box-plot پارامتر CL	۳۸
نمودار ۳-۴ چولگی داده های CL مشاهداتی در ایستگاه دهکده سفید	۳۹
نمودار ۴-۴ چولگی داده های CL مشاهداتی در ایستگاه درودزن	۳۹
نمودار ۵-۴ چولگی داده های CL مشاهداتی در ایستگاه حسن آباد	۴۰
نمودار ۶-۴ چولگی داده های CL مشاهداتی در ایستگاه چمریز	۴۰
نمودار ۷-۴ چولگی داده های CL مشاهداتی در ایستگاه پلخان	۴۱
نمودار ۸-۴ چولگی داده های دبی مشاهداتی در ایستگاه پلخان	۴۱
نمودار ۹-۴ چولگی داده های دبی مشاهداتی در ایستگاه چمریز	۴۲
نمودار ۱۰-۴ چولگی داده های دبی مشاهداتی در ایستگاه درودزن	۴۲
نمودار ۱۱-۴ چولگی داده های دبی مشاهداتی در ایستگاه دهکده سفید	۴۳
نمودار ۱۲-۴ پراکندگی داده های CL و معادله روش OLS در ایستگاه های OLS حسن آباد-پل خان	۴۴
نمودار ۱۳-۴ پراکندگی داده های CL و معادله روش OLS در ایستگاه های حسن آباد-درودزن	۴۵
نمودار ۱۴-۴ پراکندگی داده های CL و معادله روش OLS در ایستگاه های درودزن-چمریز	۴۶
نمودار ۱۵-۴ پراکندگی داده های CL و معادله روش OLS در ایستگاه های دهکده سفید-چمریز	۴۷
نمودار ۱۶-۴ پراکندگی داده های دبی و معادله روش OLS در ایستگاه های OLS دهکده سفید-چمریز	۴۸
نمودار ۱۷-۴ پراکندگی داده های دبی و معادله روش OLS در ایستگاه های OLS درودزن-چمریز	۴۹
نمودار ۱۸-۴ پراکندگی داده های دبی و معادله روش OLS در ایستگاه های OLS درودزن-پل خان	۵۰
نمودار ۱۹-۴ پراکندگی داده های دبی و معادله روش OLS در ایستگاه های OLS حسن آباد-پل خان	۵۱

نمودار ۴-۲۰ پراکندگی داده های دبی و معادله روش OLS در ایستگاه های OLS حسن آباد-  
درودزن ..... ۵۲

## فهرست شکل ها

<u>عنوان</u>	<u>شماره صفحه</u>
شکل (۱-۲) نمایی از موقعیت حوزه آبریز رودخانه کر و موقعیت جغرافیایی ایستگاههای آبسنجی (مطالعات پایه، ۱۳۷۹).....	۱۴
شکل (۲-۲) شماتیک شبکه آبراهه ها و سرشاخه های رودخانه کر.....	۱۵

## چکیده :

داده های ناقص دبی یکی از مشکلات متخصصان و طراحان پروژه های منابع آب است و باعث بروز خطا در نتایج مطالعات طرح ها شده و اجرای پروژه ها را دچار مشکل می نماید، در مناطقی که ایستگاههای هیدرومتری از نظر تعداد محدود می باشند این مسأله حادثتر می باشد، بنابراین لازم است این نواقص آماری به طریقی برطرف گردد. ساده ترین روش بازسازی Ordinary least squares regression (OLS) (رگرسیون به روش مینیمم مربعات) می باشد که در این برآورد صدک به طور نرمال توزیع شده است و روش line of organic Correlation (LOC) (خط همبستگی) و Kendall-Theil robust line (KTRL) (خط kendal-Theil) نیز از این روش پیروی می کند. یکی از مناسبترین روشها robust line of organic correlation (RLOC) (خط همبستگی -بنیانی) است، که روشی جدید می باشد و در کاهش میزان اربیی در برآورد واریانس اثر زیادی دارد. با اینکه RLOC شبیه LOC میباشد ولی بر سه روش دیگر برتری دارد. برای استفاده از روش های معرفی شده و باز سازی داده ها بعد از مرتب کردن داده ها بر اساس تاریخ و انتخاب داده های مشترک بر اساس ماه، مقادیر آن ها از طریق معادلات حاصله برآورد گردید. هدف این تحقیق مقایسه روش های بازسازی OLS، KTRL، LOC، RLOC با استفاده از شاخص های آماری ریشه دوم میانگین مربعات خطا (RMSE)، میانگین خطای مطلق (MAE) و میانگین خطای (MBE) bias جهت انتخاب بهترین روش برای بازسازی داده های دبی و کلر (CL) رودخانه کر می باشد. طبق نتایج بدست آمده مقدار RMSE, MAE, MBE روش OLS از دیگر روشها کمتر بوده و به عبارت دیگر به علت پایین بودن خطا نسبت به دیگر روشها بهترین روش برای بازسازی داده های دبی و CL رودخانه کر شناخته شد. و بعد از آن به ترتیب روش های KTRL و LOC و RLOC دارای کمترین خطا بودند.

**کلمات کلیدی:** بازسازی داده های مفقود، رگرسیون به روش مینیمم مربعات، خط همبستگی، میانگین مربعات خطا، خط همبستگی بنیانی.



# فصل اول: کلیات تحقیق

## ۱-۱ مقدمه:

اساس و پایه مطالعات هیدرولوژی، داده های آماری مورد قبول می باشد با توجه به خلأهای گسسته و پیوسته در اغلب داده های هیدرولوژی مانند دبی رودخانه ها که به دلایلی مانند ثبت نکردن آمار، حذف آمار غلط و خرابی یا از بین رفتن دستگاه های اندازه گیری اتفاق می افتد، تخمین و برآورد این داده ها لازم و ضروری است، دسترسی به داده های کافی و دقیق از یک سو موجب کوتاه تر شدن مدت مطالعات و از سوی دیگر موجب برآورد دقیق تر پارامترهای هدف و کاهش هزینه های اجرایی و خسارات بعدی ناشی از اجرای طرح ها می گردد. (لوک زاده، ۲۰۰۴)

در بسیاری از موارد، مدیریت منابع آب مستلزم استفاده از داده های هیدرولوژیک یا کیفیت آب متفاوتی به منظور شبیه سازی نتایج تصمیمات می باشد (هرش، ۱۹۸۲) اگر چه داده های مشاهداتی در دسترس برای بسیاری از جریان ها یا آنقدر کوتاهند که شامل دامنه کافی از شرایط هیدرولوژیک یا کیفیت آب نمی شوند یا داده های دوره ای در دسترس نمی باشد. (الی و برنز، ۱۹۸۳)

یک راه حل برای این مساله این است که از داده های ایستگاه های نزدیک با داده های مشاهداتی طولانی مدت در دسترس استفاده شود. (هرش، ۱۹۸۲؛ الی و برنز، ۱۹۸۳؛ ووگل و استدنیگر، ۱۹۸۵).

این روش را می توان با بسط داده های مشاهداتی کیفیت آب و هیدرولوژیک پیشینه به موقع از طریق همبستگی میان این داده های مشاهداتی در سایت مورد بررسی و داده های مشاهداتی جاری در سایتی نزدیک صورت داد و معمولاً با استفاده از رگرسیون مربع (OLS) انجام می شود، به طور معمول OLS به منظور ساماندهی اطلاعات در مورد متغیرهای کیفیت آب کوتاه سنج به کار می رود. (هارمون سیوگلو و یوجوش، ۱۹۸۶، ۱۹۸۷؛ هارمون سیوگلو و همکاران، ۱۹۹۹ (رویسون و همکاران، ۲۰۰۴؛ خلیل و اوردا، ۲۰۰۹)

از نظر تعریفی داده های کیفیت آب غیر منفی هستند و دارای ویژگی های خاصی از قبیل حضور داده های پرت، انحراف مثبت، پراکندگی غیر معمول داده ها، داده های مشاهداتی سانسور شده (مثل تمرکز زیر حد آشکار)، الگوی فصلی و خود همبستگی می باشد. داده های پرت و انحراف مثبت در داده های آب رایج تر هستند، به دلیل حضور داده های پرت و انحراف مثبت، داده های کیفیت آب اغلب شکلی مشابه به توزیع نرمال لگاریتمی دارند. (لتنمیر، ۱۹۸۸؛ بری من و همکاران، ۱۹۸۸)

دو نقص عمده در استفاده از Ordinary least squares regression (OLS) برای داده های کیفیت آب وجود دارد، اول اینکه، در حضور داده های پرت پایدار نیست، حضور داده های پرت به طور عمده بر

شیب خط در OLS تاثیر می گذارد (نویت و تام، ۱۹۸۹؛ گرانتو، ۲۰۰۶). دوم اینکه واریانس داده های بسط داده شده برآوردی از واریانس واقعی است که گرایشی منفی دارد. (هرش، ۱۹۸۲؛ الی و برنز، ۱۹۸۳؛ موگ و ویتینگ ۱۹۹۹؛ هلسل و هرش، ۲۰۰۲؛ خلیلی و همکاران، ۲۰۱۰؛ کوتسویانیس و لانگوسیس، ۲۰۱۱)

اگر روش مورد استفاده گرایشی در تخمین مقادیر حدی داشته باشد، منجر به گرایش به بیش تخمینی می گردد (هرش، ۱۹۸۲) برای کیفیت آب، مقادیر حدی و صدک ها برای ارزیابی سازگاری با استانداردها یا حدود مجاز اهمیت دارد. (خلیلی و همکاران، ۲۰۱۰).

در متون گذشته چندین روش رگرسیون پایدار مشابه OLS وجود دارد با این مزیت که در حضور داده های پرت پایدارترند (مثل تخمین Huber کمینه متوسط مربع، کمینه انحراف کل، رگرسیون وینسوری<sup>۱</sup> و مربع های کمینه اصلاحی)، به طور کلی روش های رگرسیون پایدار برای شرایطی توسعه یافته اند که توزیع های متقارن خطا به دلیل وجود داده های پرت در داده های مشاهداتی، سبب تاثیر گذاری بر مقادیر حدی می شوند. (دیتز، ۱۹۸۷؛ نویت و تام، ۱۹۸۹) بر خلاف این، در روش های غیر پارامتریک مثل رگرسیون مونوتیک<sup>۲</sup> و خط پایدار (khendall-theil)، تخمین پارامترها بدون در نظر گرفتن توزیع خاصی صورت می پذیرد.

(نویت و تام، ۱۹۸۹) رفتار روش های رگرسیون های پایدار و رگرسیون غیر پارامتریک را با OLS با توجه به حضور داده های پرت و انحراف از حالت نرمال مقایسه کردند. نتایج نشان داد خط پایدار کندال-تیل (KTRL) بسیار شبیه OLS می باشد، با این تفاوت که در حضور داده های پرت و انحراف از حالت نرمال روش KTRL برتر است، اگر چه نتایج نشان داد که حداقل انحراف مطلق - Least Absolute

(LAD) Deviation در حضور داده های پرت بسیار بهتر از KTRL عمل می کند، اما KTRL از عملکردی مشابه با رگرسیون LAD برخوردار بود. در انحراف از حالت نرمال نیز، هیچ تخمین گری بهتر از KTRL عمل نکرد، بنابراین نویت و تام (۱۹۸۹) نتیجه گرفتند که روش KTRL تخمین مستحکمتری از پارامترها در حضور داده های پرت و یا انحراف از حالت نرمال فراهم می کند.

هلسل و هرش، (۲۰۰۲) نشان دادند عملکرد خط پایدار KTRL مشابه OLS است، با این مزیت که در حضور داده های پرت و یا انحراف از حالت نرمال پایداری باشد. KTRL نیز با داده های نرمال تقریباً کارایی OLS را دارد ولی در انحراف از حالت نرمال کارا تر است.

---

۱ Vinestory

۲ Monotic

برای موردی که داده ها الگوی خطی، واریانس Homoscedastic و حالت نرمال نقاط را نشان می دهند، نتایج KTRL و OLS تقریباً مشابه می باشد (هرش و همکاران، ۱۹۹۱) وقتی داده های پرت وجود دارد، KTRL خطی کارایی بیشتری از OLS نشان می دهد، (هرش و همکاران، ۱۹۹۱؛ هلسل و هرش، ۲۰۰۲) اگر چه به مانند OLS، KTRL نیز واریانس داده های بسط داده شده را کمتر از واقعیت برآورد می کند، KTRL را نه تنها برای بسط داده ها بلکه برای ارزیابی روند نیز به کار می برند. (مانند البک، ۲۰۰۳؛ گراتانو، ۲۰۰۶؛ اولسون و همکاران، ۲۰۱۰؛ دری و همکاران، ۲۰۱۱).

کرتیسکی و منکل (۱۹۶۸) برای اولین بار خط همبستگی ارگانیک (LOC) را مورد استفاده قرار دادند. مهم ترین مزیت LOC این است که قادر به حفظ خصوصیات تغییرات مشاهداتی در داده ها است، اگر چه در حضور داده های پرت پایدار نیست، (هلسل و هرش، ۲۰۰۲). چندین پژوهش نیز از LOC برای بسط داده های جریان مانند هرش (۱۹۸۲)، هرش و همکاران (۱۹۹۱) جیا و کولور (۲۰۰۶)، ریو و همکاران (۲۰۱۰)، تخمین مقادیر بارندگی مفقود مانند رازعی و همکاران (۲۰۰۹) و (۲۰۱۱) و برای بسط داده های کیفیت آب (مانند خلیل و همکاران، ۲۰۱۱، ۲۰۱۰) استفاده کرده اند.

هدف عمده تحقیق حاضر ارزیابی مناسب بودن ۴ روش بسط داده های کیفیت آب در مناطقی که با کمبود ایستگاه روبرو هستند، می باشد. روش های مورد بررسی شامل OLS، KTRL، LOC و RLOC اصلاح شده است با توجه به اینکه روش LOC از عملکرد بهتری در حفظ خصوصیات تغییرات مشاهداتی در داده های بسط داده شده برخوردار است. از مدل اصلاح شده RLOC تحت عنوان خط پایدار همبستگی ارگانیک استفاده می شود.

#### ۲-۱ فرضیه ها :

۱- روش همبستگی بنیانی در مقایسه با روشهای LOC (خط همبستگی) و OLS (رگرسیون به روش مینیمم مربعات) و KTRL (خط Kenda-Theil) مناسب ترین روش جهت بازسازی داده های مفقود است.

۲- به دلیل وجود داده های پرت روش KTRL عملکرد بهتری نسبت به OLS دارد.

۳- عملکرد KTRL و OLS شبیه به هم می باشد.

۴- در رگرسیون گیری روش RLOC نیاز به نمونه بزرگتر نسبت به روش LOC دارد.

۵- LOC قادر به حفظ تنوع و گستردگی داده ها می باشد.

۶- در بازسازی صدک های بالا استفاده از RLOC و در صدک های پایین استفاده از LOC ارجح است.

#### ۳-۱ اهداف تحقیق:

۱- استفاده از روشی بهینه از بین روشهای مذکور جهت بازسازی داده های مفقود