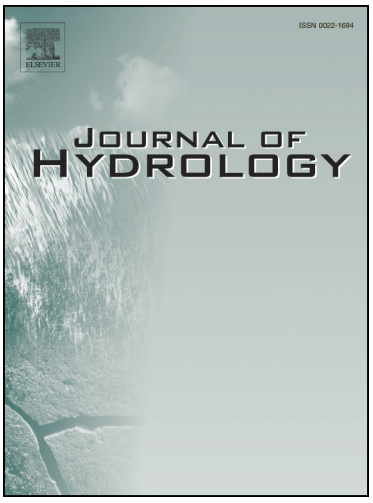
Journal Pre-proofs



Using spacetime geostatistical analysis to improve precipitation isoscape interpolation in Australia

Candida M. Duff, Jagoda Crawford, Ryan H.L. Ip, Zhenquan Li, Catherine E. Hughes, Carol V. Tadros

PII: S0022-1694(24)01898-5
DOI: https://doi.org/10.1016/j.jhydrol.2024.132502
Reference: HYDROL 132502

To appear in: *Journal of Hydrology*

Received Date: 23 December 2023
Revised Date: 20 November 2024
Accepted Date: 25 November 2024

Please cite this article as: Duff, C.M., Crawford, J., Ip, R.H.L., Li, Z., Hughes, C.E., Tadros, C.V., Using spacetime geostatistical analysis to improve precipitation isoscape interpolation in Australia, *Journal of Hydrology* (2024), doi: https://doi.org/10.1016/j.jhydrol.2024.132502

Title: Using spacetime geostatistical analysis to improve precipitation isoscape interpolation in Australia.

Author names and affiliations:

Candida M. Duff [ab]*, Jagoda Crawford [c], Ryan H. L. Ip [ad], Zhenquan Li [e], Catherine E. Hughes [c,f], Carol V. Tadros[cf]

a. Charles Sturt University, Wagga Wagga NSW 2678, Australia.
b. Research Institute for the Environment and Livelihoods, Charles Darwin University, NT, Australia
c. ANSTO, Locked Bag 2001 Kirrawee, DC NSW 2232 Australia.
d. Department of Mathematical Sciences, Auckland University of Technology, Auckland, New Zealand
e. Charles Sturt University, Albury NSW 2640, Australia
f. School of Biological, Earth and Environmental Sciences, UNSW Sydney, NSW 2052, Australia

***Corresponding author:** Candida M. Duff. Email: candy.duff@cdu.edu

CRediT author statement:

Candida M. Duff: Conceptualisation, Methodology, Software, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization. **Jagoda Crawford**: Validation, Writing - Review & Editing, Supervision. **Ryan H. L. Ip:** Conceptualisation, Software, Writing - Review & Editing**,** Supervision. **Zhenquan Li:** Software, Writing - Review & Editing**,** Supervision. **Catherine E. Hughes:** Conceptualisation, Writing - Review & Editing**,** Supervision. **Carol V. Tadros:** Data Curation, Writing - Review & Editing**.**

Highlights

- Algorithm for universal spacetime cokriging was developed for R.
- Universal spacetime kriging and cokriging improved interpolation.
- Climatic covariates more influential than purely spatial covariates.

Abstract

Isoscapes are an invaluable tool for understanding the spatial patterning of stable precipitation isotopes around the world. Researchers have used many different interpolation methods to create isoscapes from simple deterministic models, using a weighted average, to stochastic methods like machine learning (ML), universal kriging (UK) and cokriging (CK). However, to the best of our knowledge, the spatio-temporal geostatistical method using universal spatio-temporal kriging (USK) and universal spatio-temporal cokriging (USCo) has not yet been applied. Additionally, there is no current function in the geostatistical packages for R programming software, which can implement USCo. This study aims to develop an interpolation method that identifies both the spatial and temporal correlation structure of isotopic signatures in data poor regions. This method will allow for interpolation on a month-to-month basis, improving temporal resolution when compared to previous long-term averages for annual or monthly isoscapes. An algorithm for USCo was developed for R. Historical stable precipitation isotope data ($\delta^{18}O$ and $\delta^{2}H$) from southeastern Australia were analysed using the spatio-temporal geostatistical method. Results from a leave-one-out analysis, compared generalised least squares regression analysis (GLS), USK and USCo. Metrics used were root mean square error (RMSE); mean absolute error (MAE); predicted residual error sum of squares (PRESS); Akaike information criterion (AIC); and Bayesian information criterion (BIC).

Both USK and USCo consistently outperformed GLS, with the best MAE for $\delta^{18}O$ being 1.75 (GLS), 1.35 (USCo) and 1.32 (USK). For $\delta^{18}O$ USK outperformed USCo (MAE 1.32 vs 1.35) whereas in models for $\delta^{2}H$ USCo outperformed USK (MAE 10.44 vs 10.79). However, these differences were small, and the instability and the long processing time required for the USCo algorithm (480 min. vs 0.35 min. for USK) made USK the preferred method. Additionally, the isoscapes generated from USCo contained several 'bullseyes', detracting from the preferred smoothing effect expected from kriging. We found that the isotopic spatial patterning was similar for adjacent months but quite different as time progressed. Climatic covariates were better predictors than purely spatial covariates like distance from the coast and altitude. This was probably due to the strong spatial patterning of climate over the study area. We identified some very depleted isotope values in June 2007, which were probably due to large precipitation events at that time. Some areas, for example, south of Sydney, were well documented giving good interpolation results. Others would benefit from more data collection. The methods and models discussed in this paper can greatly improve the temporal resolution of isoscape models for precipitation isotopes. This is particularly useful in regions where isotope ratio data is sparse, and direct observations are available to enable the statistical filling of these spatial gaps.

Key Words: isoscapes, kriging, cokriging, precipitation, precipitation isotopes.

1 Introduction

Bowen (2010b) states that one of the most powerful examples of the global-scale patterning of stable water isotopes ($\delta^{18}O$ and $\delta^{2}H$) is their partitioning during phase-change reactions in the water cycle. Analysis of these patterns has been of interest since they were revealed in studies conducted in the 1950s (e.g., Dansgaard, 1954; Dansgaard, 1964; Friedman, 1953). In response, the International Atomic Energy Agency (IAEA) and the World Meteorological Organization (WMO) have developed the Global Network of Isotopes in Precipitation (GNIP), a worldwide monitoring network of oxygen and hydrogen isotopes in precipitation (IAEA/WMO, 2022). GNIP and other regional monitoring programs have been crucial in understanding the factors that affect the $\delta^{18}O$ and $\delta^{2}H$ values of precipitation (Araguás-Araguás et al., 2000; Crawford et al., 2016; Jiao et al., 2019; Kaseke et al., 2016; Liu et al., 2010; Matiatos & Wassenaar, 2019; Munksgaard et al., 2019; Wang et al.,

2022; Xu et al., 2021). This has led to the development of isoscapes (thematic maps), which depict the $\delta^{18}O$ and $\delta^{2}H$ values of precipitation over a geographical area (e.g., Allen et al., 2018; Bowen et al., 2022; Hollins et al., 2018; Kaseke et al., 2016; Nlend et al., 2023; Wang et al., 2022).

Important factors that affect the $\delta^{18}O$ and $\delta^{2}H$ values of precipitation, include the conditions at the site of evaporation (i.e., sea surface temperature, relative humidity, and wind speed), mixing along the vapour trajectory), and temperature at the point of condensation (Dansgaard, 1964). Additionally, the local climate has a significant impact because as the rain falls, the raindrops exchange moisture with the surrounding vapour, and sub-cloud evaporation occurs (Crawford et al., 2016). As distance from the coast and elevation increases, continental (Liu et al., 2010) altitudinal (e.g., Araguás-Araguás et al., 2000; Jiao et al., 2019; Wang et al., 2022; Xu et al., 2021), and slope (i.e., elevation rate of change, Nlend et al., 2023) effects are observed. Table 1 summarises some of the climatic factors that have been used in the interpolation of $\delta^{18}O$ and $\delta^{2}H$ values.

[Include Table 1 here]

There are many different approaches for modelling isoscapes. For example, deterministic methods such as inverse distance weighting and spline interpolation (Bowen & Revenaugh, 2003; Xu et al., 2021). Stochastic regression models, which consider explanatory variables, have been used extensively (e.g., Birkel et al., 2018; Bowen & Revenaugh, 2003; Hollins et al., 2018; Terzer-Wassmuth et al., 2021; Terzer et al., 2013; Wassenaar et al., 2009). More recently, stochastic *geostatistical* methods such as ordinary or universal kriging and cokriging, have been used (Bedaso & Wu, 2021; Belkhiri et al., 2020; Bowen et al., 2022; Hatvani et al., 2020; Kaseke et al., 2016; Rostami et al., 2020; Xu et al., 2021) and increasingly, geostatistical methods are being augmented with machine learning (e.g., Erdélyi et al., 2023). Deterministic methods assume known mean rates. However, this is often not the case with $\delta^{18}O$ and $\delta^{2}H$ values. Whilst spatial geostatistical models may improve interpolation when compared to simple regression models (Cressie, 1993), and further improvement may be achieved with machine learning (e.g., Erdélyi et al., 2023; Gribov & Krivoruchko, 2020), these approaches do not utilise the temporal information contained within the data. Many authors (Cressie & Wikle, 2011; González et al., 2016; Schabenberger & Gotway, 2017; Subba Rao & Terdik, 2017) argue that there is a need to address the *time* component of spacetime processes.

Recently interest in the temporal analysis of isotopes has strengthened (e.g., Matiatos & Wassenaar, 2019; Nlend et al., 2023; Wang et al., 2022). Temporal precipitation isoscapes (sometimes called 'timescapes') have been created, using machine learning or regression (Delavau et al., 2015; Erdélyi et al., 2023; Nelson et al., 2021; Wang et al., 2022). However, these studies had considerably more extensive datasets than were available for this study. By using the spatio-temporal geostatistical approach, we have been able to produce useful temporal precipitation isoscapes, that illustrate the temporal correlation structure, with a comparatively small amount of data. Additionally, we have published our R code to enable other researchers to apply this method to their own datasets and expand its application.In places where the unpredictable climate means that $\delta^{18}O$ and $\delta^{2}H$ values vary widely from year to year, monthly isoscapes that rely on annual averages of climatic variables for isotopic values are expected to be inaccurate, as demonstrated by Hollins et al. (2018). By using monthly data from each year, and implementing the spacetime (ST) geostatistical approach, interpolation accuracy can be greatly improved. Additionally, the implementation of the universal ST cokriging (USCo) algorithm in R (R Core Team, 2022), may improve interpolation, as it uses the strong correlation between $\delta^{18}O$ and $\delta^{2}H$ to boost performance.

To test the ST geostatistical approach, a case study has been created for southeastern Australia and includes an area with topographic contrast, unpredictable climate, a sparce dataset, and a range of climate zones. The choice of this heterogeneous area will demonstrate how the method can be used

globally. Specifically, the results of the USK and USCo will be compared to each other and to linear models calculated using generalised least squares (GLS) using the following metrics to measure the model performance: root mean square error (RMSE), mean absolute error (MAE), predicted residual error sum of squares (PRESS), Akaike information criterion (AIC) and Bayesian information criterion (BIC).

Using the ST geostatistical approach described herein will allow researchers to predict more accurately the $\delta^{18}O$ and $\delta^{2}H$ values in areas, and at times, that have not been sampled when compared to ordinary kriging and other regression methods. This study aims to 1) determine whether $\delta^{18}O$ and $\delta^{2}H$ values exhibit both a spatial and temporal correlation, and if so, describe the spatio-temporal structure; 2) compare the spacetime geostatistical approach to other regression approaches; 3) develop an algorithm in R (R Core Team, 2022) for universal spacetime cokriging (USCo); and 4) quantify the uncertainties to determine where more data could be collected in future studies. Future spacetime precipitation based isoscape models created in this manner, can be used at local and global scales in hydrological, ecological, provenance, and forensics applications.

2 Site description and methods

2.1 Study area

Apart from Antarctica, Australia is the driest and flattest continent on Earth, with much of the country experiencing hot summers and mild or cold winters. The study area is situated in southeastern Australia covering an area of $19.8 \times 10^8\ km^2$(Figure 1). It includes the states of Victoria, New South Wales, and an area of southern Queensland south of latitude 26° S. Topographic contrast is provided along the east coast, from around Melbourne to north of Brisbane, by the Great Dividing Range. The winters are cold, with warm to mild summers along the coast, and hot dry summers inland. There is an arid area in the northwest and a range of precipitation zones closer to the coast (see Figure 1 and Supplementary materials Section F). Most of the collection sites are situated in the wet summer and low winter or uniform precipitation zones, with no collection sites in the arid area and only three sites in the marked wet winter and dry summer zone. This lack of representative sites led to some extrapolation for these areas.

There were 32 collection sites in the study area, with most sites situated closer to the coast, and just five sites west of longitude 146° E and north of latitude 37° S. The greatest distance between sites was Melbourne to Brisbane (1380 km) and greatest distance from the nearest coast was Charleville (554 km).

[Include Figure 1 here]

2.2 Data collection

Australian precipitation isotope data were obtained from the GNIP (IAEA/WMO, 2022) and other previous studies by ANSTO, CSIRO and Geoscience Australia (i.e., Crawford et al., 2016; Crosbie et al., 2012; Hawkins et al., 2022; Hollins et al., 2018; Hughes & Crawford, 2013; Ransley et al., 2015; Tadros et al., 2022) (see Table 2). The collection and analysis methods of these data are presented in those studies.

[Include Table 2 here]

Climate data for the entire study area (Table 3) were obtained from Bureau of Meteorology (BoM) (2023b). Altitude/elevation (meters above sea level; MASL) and slope (%) data were extracted from the Digital Elevation Model (DEM) of Australia (Geoscience Australia, 2018) using a 0.1-degree grid.

Slope, which refers to rate of change in elevation, has been considered an important factor, especially for diverse topographies (Nlend et al., 2023), therefore, it was included in the analysis. Using the method described by Terzer-Wassmuth et al. (2021), distance from the nearest coast (km) was calculated employing the R packages **sf** (Pebesma, 2018), **rnaturalearth** (South, 2017a) and **rnaturalearthdata** (South, 2017b).

[Include Table 3 here]

Aside from the spatial covariates (e.g., distance, slope, latitude, altitude), many climatic covariates have been found to help explain the ST variation of $\delta^{18}O$ and $\delta^{2}H$ values. Some of these are summarised in Table 1. In this study, average dewpoint (°C month^{-1}), average areal actual evapotranspiration (mm month^{-1}), total precipitation (mm month^{-1}), average relative humidity (% month^{-1}), maximum temperature (°C month^{-1}) and average vapour pressure (hPa month-1) were considered. Most of the climatic data was collected and calculated from actual monthly 2007-2008 observations available from BoM (2023b). Composite data for average areal actual evapotranspiration (mm month^{-1}), which are based on the standard 30-year period 1961-1990 from BoM, (2023b), were used in this study, as data for individual years were not available. As latent heat flux, synoptic type and wind speed are more suited to studies with daily observations, they were not considered. In Australia, longwave radiation data is not readily available from the Australian Bureau of Meteorology (2023a) so it was decided that it should be omitted as well.

Dewpoint (DP) was derived from the Vapour Pressure (VP) equation (1) published by BoM (2023a) and used to calculate relative humidity (RH). The equation for RH (2), was derived from that published by Lawrence (2005).

$$\mathrm{VP} = \exp\left(1.8096 + \frac{17.269425 * \mathrm{DP}}{237.3 + \mathrm{DP}}\right) \quad (1)$$

$$RH = 100\left(\exp\left(\phi \cdot \left(\frac{DP\cdot(\varphi+T) - T\cdot DP - T\cdot\varphi}{\varphi\cdot(\varphi+T)}\right)\right)\right)^{\frac{\varphi}{DP+\varphi}} \quad (2)$$

Where DP is the dewpoint, T is the temperature, φ and $\phi$are the Magnus coefficients recommended by Alduchov and Eskridge (1996).

2.3 Data preparation

Monthly precipitation isotope data were collated and, along with the covariates, placed into a single datafile. As the intention was to produce monthly isoscapes and a day must be specified for processing, the collection date was set to the 15th of each month. This allowed for roughly even time lags. The dates were put into the format “yyyy-mm-dd”. The mean monthly climate data was available at either a 0.1° or 0.05° resolution, therefore, the precision of the dataset’s longitudes and latitudes was reduced to two decimal places. This new ‘clean’ dataset was saved for use in the analysis.

A 0.1° rectangular grid was created covering the entire study area and then clipped to the shape of the coast using the method described by Neto (2014). The grid was populated with the covariates and sorted into date order. Grid points with monthly precipitation equal to zero were removed, avoiding illogical interpolation of values where no precipitation had fallen.

2.4 Linear model analysis

A linear model analysis was performed to find the trend in the data and analyse how each covariate affects the $\delta^{18}O$ or $\delta^{2}H$ value. To avoid issues of multicollinearity, a correlation test was performed on all predictor covariates. Covariates that were highly correlated, i.e., temperature with RH, evapotranspiration, and VP, were separated into subgroups and analysed separately. As the $\delta^{18}O$ or $\delta^{2}H$ values of precipitation are known to be very noisy three methods of analysis were considered.

- MOA1 required analysis to be performed on each individual year.
- MOA2 used the combined dataset (2007-2008) to perform the analysis.
- MOA3 was a hybrid method, where a linear analysis was performed on the full dataset to establish the best covariates, but the geostatistical analysis (i.e., the USK and USCo) was performed on the individual years separately.

The linear model analysis had four stages, first a stepwise analysis was performed to identify the covariates for the models, then the models were manually checked to ensure all coefficients had a p-value of < 0.05. The variance inflation factor (VIF) for each covariate was checked for multicollinearity. Covariates with VIF factors > 2.0 were removed. To determine the presence of autocorrelation, a Durbin-Watson (DW) test was performed on each model. The DW statistic ranges from zero to four; values close to two indicate no autocorrelation, however, as the DW statistic was significantly less than two, GLS models were created for comparison to ST geostatistical models. Further details of the linear model analysis can be found in Section A of the Supplementary Materials.

2.5 Universal spatio-temporal kriging

Ordinary spatio-temporal kriging is an interpolation method that predicts the value in an unobserved location at an unobserved time through a weighted linear combination of surrounding spatio-temporal values (Cressie & Wikle, 2011). It uses the ST variogram to model the spatio-temporal correlation structure of the data. USK extends ordinary spatio-temporal kriging in that it includes a trend which is represented by a linear model (Cressie & Wikle, 2011). It is defined as:

$$\hat{Z}(\boldsymbol{s},t) = \boldsymbol{X}\cdot\boldsymbol{\beta} + \boldsymbol{\lambda}\cdot\boldsymbol{Z}(\boldsymbol{s},t), \sum_{i=1}^{N}\lambda_i = 0 \quad (3)$$

Where $\hat{Z}(\boldsymbol{s},t)$ is the predictions either for $\delta^{18}\hat{O}$ or $\delta^{2}\hat{H}$ at coordinates $\boldsymbol{s_0}$ at time t_0; $\boldsymbol{X}\cdot\boldsymbol{\beta}$ is the linear model representing the trend ($\boldsymbol{X}$ represents the factors and $\boldsymbol{\beta}$ represents to coefficients); N is the number of observations; and $\boldsymbol{\lambda}$ is the vector of weights given to each of the observations in the vector $\boldsymbol{Z}(\boldsymbol{s},t)$.

For the analysis methods MOA1 and MOA3, USK was performed using the **gstat** package (Graler et al., 2016; Pebesma & Bivand, 2005a) in R (R Core Team, 2023) first on 2007 data (using the identified covariates) and then on the 2008 data. The results were concatenated for the isoscape. MOA2 applied USK to the complete dataset (2007-2008) using the covariates identified from the combined 2007-2008 dataset. USK was performed as detailed below:

Following the method described by Wikle et al. (2019), a spatial-point object was created, from the unique coordinates in the dataset. The coordinate reference system (CRS) was set to WGS84, using the **sp** package (Bivand et al., 2013; Pebesma & Bivand, 2005b) and R (R Core Team, 2023). The dataset did not contain data at every point for every month, so an ST data-frame for data with partial ST grids (STSDF) was created using the **spacetime** package (Bivand et al., 2013; Pebesma,

2012). STSDF objects are part of a suite of objects that hold the data for creating ST variograms in the **gstat** package (for more details see Graler et al., 2016; Pebesma, 2023; Pebesma, 2014).

The wide range combined with the left skewness of the δ^2H data produced errors in processing, therefore, δ^2H was divided by eight to narrow the range. Eight was chosen as a divisor because the global meteoric water line (GMWL) is defined as $\delta^2H = 8 \cdot \delta^{18}O + 10$ (Bowen, 2010b). Once interpolation was complete, the results were then back transformed by multiplying by eight.

STSDF objects require a $(2 \times n)$ index matrix, containing the spatial and temporal indices which must be sorted into date order. The time index for the 2007-2008 combined dataset (MOA2) was calculated using the formula $id_t = M + 12 * (Y - 2007)$. Where M is the number of the month and Y is the year $[2007,2008]$. The month number was used as the temporal index for MOA1 and MOA3. The spatial index $(i = 1, 2, \ldots n)$ where n is the number of spatial points, was generated using the unique coordinates of the dataset. Finally, an STSDF object was created containing the index matrix, the observed values for $\delta^{18}O$ and δ^2H, and the covariates identified in the linear model analysis (see Section 3.1). The STSDF object was used to create empirical variograms following the methods described by Graler et al. (2016) and Wikle et al. (2019).

The temporal and spatial ranges of the variogram models were obtained by viewing the final row of the empirical variogram output. The plots of the empirical variograms indicated the approximate values for nugget and partial sill, which is an approximate measure of half of the actual sill. The range is the maximum spatial and temporal range of the data (in kilometres and days), and the nugget describes how the data behaves near zero. A detailed explanation of variograms and variogram models can be found in (Cressie, 1993; Cressie & Wikle, 2011; Wikle et al., 2019).

The empirical variograms were analysed and, following the method described by Graler et al. (2016), were fitted to five different model classes (separable, product-sum, metric, sum-metric and simple sum-metric , see Supplementary Materials Table S6). All possible combinations of exponential and spherical variograms, and spatial, temporal, and joint variograms were used, creating a total of 32 models in all (see Supplementary Materials Figure S4). The spatial and joint variograms were given an initial spatial range of 527 km and nugget of zero. The partial sills were set to the value indicated by the associated empirical variogram. The temporal variograms were created with range of 334 days for 2007, 335 days for 2008, in MOA1 and MOA3, and 700 days for MOA2. Once all the empirical variograms were fitted to the models, the weighted MSEs (wMSE) were compared, and the models with the lowest wMSE were used for further analysis.

A grid was created using the method described by Neto (2014) incorporating a shape file, that bounded the study area, that was created using ArcGIS Pro version 2.7.0 (ESRI, 2020). The projection was set to WGS84 and resolution set to 0.1°. The midpoint of each grid cell was used as a point of reference for the grid data. The covariate data, dates and coordinates were added to the grid which was sorted into date order and used to create an STSDF object as described above. For MOA1 and MOA3 the grid data file was split into annual groups. The krigeST method from the **gstat** package (Graler et al., 2016; Pebesma & Bivand, 2005a) was used to perform the USK using the two STSDF objects and the variogram model as parameters, results of which were used to create the isoscapes.

2.6 Universal spatio-temporal cokriging

USCo extends USK by including an additional covariable into the kriging algorithm (Hu & Shu, 2019; Wackernagel, 2010). It should not be confused with USK which uses a linear model to represent the trend in the data. As there is currently no package available in R (R Core Team, 2023) for USCo, we implemented the method described by Hu and Shu (2019) and cokriged the standardised residuals (4) from the linear models derived in Section 2.4. We then back transformed the standardised

residuals as described in Equation (5) and used the predictions to create monthly isoscapes. The complete algorithm for USCo can be found in Section C of the Supplementary Materials.

Many authors have recommended standardising the data before performing cokriging (see Hu & Shu, 2019; Rossiter, 2018). Therefore, the residuals were standardised using the following method:

$$Z^{\#}_{res}(\boldsymbol{s},t) = \frac{Z_{res}(\boldsymbol{s},t) - mean(Z_{res}(\boldsymbol{s},t))}{sd(Z_{res}(\boldsymbol{s},t))} \quad (4)$$

Where $Z^{\#}_{res}(\boldsymbol{s},t)$ are the standardised residuals at coordinates $\boldsymbol{s}$ at time t, and $Z_{res}(\cdot)$ are the observed residuals for either $\delta^{18}O$ or $\delta^{2}H$. Equation (5) was used to back-transform the prediction results.

$$Z_{res}(\boldsymbol{s},t) = Z^{\#}_{res}(\boldsymbol{s},t) \times sd(Z_{res}(\boldsymbol{s},t)) + mean(Z_{res}(\boldsymbol{s},t)) \quad (5)$$

2.7 Model validation

Wikle et al. (2019) suggest using leave-one-out analysis to validate interpolation models. Additionally, by estimating the kriging variance at each point of the isoscape, the accuracy of the interpolation can be judged (Thompson et al., 2012). To assess the model performance and select the best method, the statistical criteria RMSE, MAE, PRESS, AIC and BIC were used.

Kutner (2005) recommends the use of the AIC, BIC and PRESS criteria for regression model selection and validation. AIC and BIC both penalise models for having large numbers of parameters and should be minimised. The PRESS metric measures how well the model predicts the observed values (Kutner, 2005). RMSE is also another measure of how well the model fits and is commonly used to assess model performance when interpolating the isotopic composition of modern meteoric precipitation (e.g.,Bowen & Revenaugh, 2003). Other authors (e.g., Allen et al., 2018) have used the MAE statistic which measures the average size of the errors within the prediction set omitting the error direction. This criterion is more resilient to outliers than RMSE and has been recommended for use in climate research (Willmott & Matsuura, 2005). Using GLS instead of ordinary least squares (OLS) is a regression method that is recommended for use with data that has some level of autocorrelation, and it is often used when analysing time series (Kutner, 2005).

A leave-one-out analysis was performed by omitting one site at a time and using USK, USCo and GLS to interpolate the missing site's observations. The kriging variance isoscapes were also created using USK and visually analysed to find where additional data should be collected to improve future isoscape production. Areas shaded in green and blue (i.e., kriging variance < 4) were considered as well documented, areas coloured yellow and red (kriging variance ≥ 4) were considered extrapolation.

The residuals from each method of analysis (MOA1, MOA2 and MOA3) were calculated and summarised to obtain their spread and used to calculate the goodness of fit statistics (RMSE, MAE, PRESS, AIC and BIC). AIC and BIC differ from the other assessment metrics, in that they add a penalty for having too large a number of covariates (Kutner, 2005), with BIC favouring a more parsimonious model. AIC and BIC are calculated using the following equations (Kutner, 2005):

$$AIC = n\,ln\left(\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2\right) - n\ln(n) + 2p$$

$$BIC = n\,ln\left(\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2\right) - n\ln(n) + p\ln(n)$$

Where Y represents the actual observation, $\hat{Y}$ is the prediction, n is the number of observations, and p is the number of parameters. All the goodness-of-fit statistics were compared to find the best interpolation method. Graphs of the predictions for each omitted site were created to illustrate the improved interpolation.

Thematic maps of the kriging variances were created from USK by selecting the computeVar option of the krigeST method within the **gstat** package (Graler et al., 2016) in R (R Core Team, 2023). This option computes the predicted kriging variances, the lower the kriging variance the higher the accuracy of the isoscapes. The kriging variance map was visually analysed to determine the accuracy of the isoscapes and to find where more data would improve future isoscapes.

2.8 Isoscape creation

As the results from the leave-one-out analysis were very close for both USK and USCo, both methods were used to interpolate the $\delta^{18}O$ and $\delta^{2}H$ values at each of the unobserved sites in the grid mentioned in Section 2.3. These predicted values were plotted using the R package **ggplot2** (Wickham & Grolemund, 2017) with the method described by Moreno and Basille (2018). The isoscapes were compared visually for discernible differences. Smoothness, bullseyes, and abrupt transitions were noted.

3 Results

3.1 Linear model analysis

The correlation tests (Table 4) indicated that vapour pressure (VP) and dewpoint (DP) where highly correlated (r = 0.99). Slope and altitude were also highly correlated (r = 0.70). Multi-correlation occurred with temperature and evapotranspiration (r = 0.42), relative humidity (RH) (r = -0.68) and VP (r = 0.63), returning a tolerance of 0.02. RH was correlated with distance from the coast (r = -0.62) and precipitation (r = 0.41), returning a tolerance of 0.55. Tolerance, which is a measure of multi-correlation, indicates how strongly variables affect each other; tolerance values > 0.25 are generally acceptable (Kutner, 2005). The solid red trendline, in Figure 2, shows a clear sinusoidal pattern indicating possible autocorrelation in the time series. The sinusoidal patterns of $\delta^{18}O$ and $\delta^{2}H$ values over time, have been noted previously (Allen et al. , 2018), and indicate the seasonal pattern of $\delta^{18}O$ and $\delta^{2}H$ values observed in precipitation; where lower values occur during the colder months and higher values during the warmer months (Dansgaard, 1964). Note, unlike simple regression analysis, geostatistical analysis assumes there is a relationship between the variables, and universal models, allow for this by using generalised least squares to fit the covariates.

After analysis, the best set of explanatory variables for the 2007 data were RH, evapotranspiration, VP, and total precipitation. The best set of explanatory variables for the 2008 data were maximum temperature and precipitation, and for the combined 2007-2008 dataset, RH, evapotranspiration, total precipitation, altitude, and latitude were the best set of variables. The reasons behind this difference in 2007 and 2008 models are unclear, however, the models excluding temperature had a similar $MSE_{\delta^{18}0}$ (2.3); and lower $AIC_{\delta^{18}O}$ (883 vs 895) and $BIC_{\delta^{18}0}$(896 vs 914) in 2007, whereas the model including temperature had a similar $MSE_{\delta^{18}0}$ (2.2); and lower $AIC_{\delta^{18}O}$ (883 vs 895) and

$BIC_{\delta^{18}0}(896\ vs\ 914)$in 2008. The models for δ^2H had similar results to $\delta^{18}O$. The best models are listed in Table 5. More detailed results are listed in Supplementary Materials Section A Linear Model Analysis.

The Durbin-Watson (DW) test returned p-values of < 0.001 for all covariates in all models, excepting latitude, which was correlated with latitude-squared, the DW statistic was always less than two [1.06, 1.19] (see Supplementary Materials Table S2). P-values of < 0.001 indicated that there was autocorrelation present in all models, and a DW-statistic less than two indicated positive autocorrelation. Additionally, the graphs of residuals vs lagged residuals (Figure 3) all confirm a strongly positive linear relationship indicating the presence of autocorrelation.

[Include Figure 2 here]

[Include Table 4 here]

3.2 Comparison of models

The weighted mean square errors (wMSE) of the fitted variograms varied significantly between the 2007 and 2008 data with wMSE values 10 times larger for 2007 data compared to 2008 data. The wMSE was lowest for the Sum Metric model so this model was used in the remainder of the analysis for both USK and USCo. Tables S8 and S9 in the supplementary materials depict a detailed analysis of the fitted variogram models and the wMSEs for each fitted variogram model type.

When using MOA1, USK had the lowest goodness of fit values: RMSE (1.78), MAE (1.32) and PRESS (1341) when predicting $\delta^{18}O$; and RMSE (14.12), MAE (10.78) and PRESS (84751) when predicting δ^2H. AIC and BIC were not calculated for MOA1 as there were a different number of parameters for each year. As MOA1 was generally a better fit than MOA2 and MOA3, isoscapes were created using MOA1, with USK and USCo for comparison (Figure 5). A kriging variance map was also created with the values generated from the USK (Figure. 5).

Graphs comparing USK, USCo and GLS using MOA1 were visually compared for differences. USCo and USK generally predicted a wider more accurate range of values than GLS (e.g., $\delta^{18}O_{GLS}$ $[-13.10‰, -0.59‰]$,$\delta^{18}O_{USCo}[-15.33‰, 4.73‰,]$, $\delta^{18}O_{USK}$ $[-14.24‰, 4.15\ ‰]$). USCo most often predicted the most depleted observations (e.g., Googong Figure 4). Despite the improved results from USK and USCo the errors for the extreme high and low values were still large (e.g., ANSTO October 2007 and Googong November 2007) highlighting the need for more research into the factors that affect isotopic values.As expected, areas with a higher density of sites (e.g., ANSTO and Googong) had more accurate results than in areas more sparsely sampled, e.g., Macquarie Marshes. The complete results of the leave-one-out analysis are listed in Table 6 and all the plots illustrating the results, are presented in Supplementary Materials Section A.3.

[Include Figure 4 here]

3.3 Universal spacetime kriging, universal spacetime cokriging models and isoscape creation

USK results for the predictions of $\delta^{18}O$, returned a range of [-14.24 ‰, 4.15 ‰] and δ^2H returned a range of [-98.73 ‰, 31.46 ‰]. These values were comparable to the observed ranges of [-16.43 ‰ , 3.82 ‰] for $\delta^{18}O_{obs}$ and [-116.00 ‰, 33.40 ‰] for δ^2H_{obs}. USCo results for the predictions of δ^{18}

$\hat{O}$, returned a range of [-15.33 ‰, 4.73 ‰] (after back transformation); and [-101.60 ‰,43.05 ‰] for $\delta^2\hat{H}$. The USCo isoscapes contained bullseyes around the observation sites, detracting from the preferred smoothing effect expected from kriging. USCo produced a wider range of results, however some predictions were outside of the observed range suggesting extrapolation. The kriging variance map produced from USK (Figure 5), indicated extrapolation (i.e., kriging variance ≥ 4) in the northwestern section of the study area and a need for more collection sites between Sydney and Brisbane. The USCo algorithm took 480 minutes to process and was prone to crashing. Therefore, the process was divided into monthly interpolation results, which were saved separately then concatenated for isoscape creation. The cokriging isoscapes can be found in Section D of the Supplementary Materials.

[Include Figure 5 here]

4 Discussion

4.1 Linear model analysis

Analysis of the correlation matrix (Table 4), proved that many of the selected covariates were highly correlated, leading to concerns of multi-correlation. As VP and DP were excessively correlated (r = 0.99), one of the covariates had to be omitted from the analysis. VP is readily available from BoM (2023a), and DP must be calculated using equation (1), therefore, VP was chosen as the preferred covariate. Additionally, the high correlation of slope and altitude (r = 0.70) indicated that only one should be chosen as an indicating variable. An examination of Figure 6(C) indicated an outliner at slope ~70 %. Removing this outlier, as presented in Figure 6 (D), indicated that slope had a slightly stronger correlation, however, as the relationship between altitude and isotopes is very well understood (Dansgaard, 1964), and altitude is widely used, it was chosen as the preferred variable.

The observed tolerance of 0.10 between temperature, VP, RH, and evapotranspiration, indicated that multi-correlation would produce unreliable results. Therefore, models omitting temperature were compared to models including temperature but omitting VP, RH and evapotranspiration. The tolerance between RH, distance, and precipitation (0.52) and between VP, RH and evapotranspiration (0.55) was well above the recommended value of 0.25 (Kutner, 2005) and therefore, was not considered an issue.

One of the assumptions of linear models is that the data and residuals are independent, however the Durbin-Watson correlation tests indicated that this was not the case, and autocorrelation was present. The graphs of residuals vs lagged residuals (Figure 3) displayed a strong positive linear correlation between residuals confirming they were not independent. This observed autocorrelation is quite common for timeseries, as events closer in time are often more similar than those further apart (Tobler, 1970). Of note was the apparent seasonal sinusoidal pattern discussed by Allen et al. (2018) (see Figure 2). Therefore, after models were suggested by using a stepwise analysis (see Supplementary Materials Section A, Part A.1 Stepwise analysis, Table S1), linear models were created using GLS rather than OLS using the covariates identified in Section 3.1 **nlme** package (Pinheiro & Bates, 2000, 2023).

[Include Figure 6 here]

4.2 Geostatistical Analysis

The empirical variograms indicated that the spatial correlation structure was stronger than the temporal correlation structure (see Supplementary materials Figure S5). There was little difference

between MOA1 and MOA3 for 2007, however the models for the 2008 dataset showed a much wider variance spread [0,10] vs [0,8] and the fitted models for MOA1 had a smoother transition over space and time than MOA3. For both 2007 and 2008 data, the sum metric model with a spherical joint variogram, had the lowest wMSE (5.5 x 10^2 and 6.67 x 10^3). Of note the wMSE for 2008 was around 10 times smaller than that of 2007. Additionally, The approximate spacetime sills of the fitted variograms for 2007 (340 km, 60 days, MOA1) were much smaller than the spacetime sills for the 2008 data (450 km, 90 days, MOA1) and the spacetime nuggets were much larger (1-2 km, 1-2 days vs 0 km, 0.1 – 0.3 days) (see Supplementary materials Table S7b). This may be due to the very depleted values observed in June 2007 and the difficultly in predicting the extreme values for $\delta^{18}O$ and $\delta^{2}H$ (see Figure 4 and Section 3.2)

4.3 Effect of covariates

An important research question, for this study, was which covariates improve the prediction accuracy for precipitation isotope levels, and how does each covariate affect the $\delta^{18}O$ and $\delta^{2}H$ values. In this section, the effects of the covariates are discussed.

Distance from the coast, which describes the continental effect, had a positive correlation with the $\delta^{18}O$ and $\delta^{2}H$ values ($\delta^{18}O_{cc} = 0.22$, $\delta^{2}H_{cc} = 0.15$). This is contrary to the results described by Liu et al. (2010). Hollins et al. (2018) emphasised that distance can be offset by other factors such as isotopically enriched moisture being recycled back into the atmosphere via evapotranspiration. Regardless, distance from the coast was found to be not statistically significant for the study area, given the other variables. In Australia, distance from the nearest coast, can be problematic as it may not indicate where the dominant precipitation is coming from Australia, precipitation can originate from many directions. From March to October, precipitation can originate from northwest cloud bands (BoM, 2023c; Holgate et al., 2020) bringing rain from the Indian Ocean as far south as the mallee region in northwestern Victoria (McIntosh et al., 2012). East coast lows affect this part of the country all year round (BoM, 2023c) bringing precipitation from the Pacific and cut-off-lows which originate from the Southern Ocean bring rain from the south (BoM, 2023c). Additionally, during the cooler months of the year, when the subtropical ridge moves further north, westerly air flow dominates the southern part of Australia. The southern region is dominated by the occurrence of cold fronts that may develop into low pressure systems. Both weather systems can bring precipitation from the Great Australian Bite (BoM, 2023c). Therefore, distance may not be a good predictor for $\delta^{18}O$ and $\delta^{2}H$ values in eastern Australia.

Degree of slope, considered by Nlend et al. (2023), was found to be highly correlated with altitude (r = 0.70), and thus, slope also had an inverse correlation with the isotopic signatures ($\delta^{18}O_{cc} = -0.21$, $\delta^{2}H_{cc} = -0.17$). This reflects the "inverse altitude effect" described by Jiao et al. (2019) and Xu et al. (2021). In Australia, slope and/or altitude can also be problematic. Apart from Antarctica, Australia is the flattest continent on Earth, even the peaks of the Great Dividing Range are not high, with the tallest mountain, Mount Kosciuszko, reaching only to 2228 meters. Therefore, both altitude and slope have a negligible effect on the inland sites. For MOA1 altitude did not have a significant effect, however when the 2007-2008 data was combined, altitude decreased the value of $\delta^{18}O$ by 0.001‰ and $\delta^{2}H$ by 0.008 ‰ for every 1 m increase in elevation.

Isotopic composition and latitude followed the bell shaped curved described by Feng et al. (2009). However, as the study area is situated in the mid-latitudes, the quadratic relationship was only marginally more significant than a linear one (see Figure 6). Nonetheless, for the combined 2007-2008 dataset the $\delta^{18}O$ the relationship was $\widehat{\delta^{18}0} = 2.26\,|lat| - 0.03\,lat^2$ and $\widehat{\delta^{2}H} = 18.97\,|lat| - 0.30\,lat^2$. Notably, when analysing each year separately, none of the spatial covariates were statistically significant. This could be due to the very depleted values observed in June 2007 or as a

result of the fact that the observed data is inherently noisy and varies widely from year to year. Alternatively, this could indicate that, in southeastern Australia, the spatial aspect of climatic variables mean they are better predictors than purely spatial variables.

The graph of total precipitation (mm) (Figure 7) indicates it has a negative relationship with the $\delta^{18}O$ and $\delta^{2}H$ values. As precipitation falls, the heavy isotopes are preferentially striped from the cloud, this results in a rainout effect and a depletion of heavy isotopes (Bowen, 2010b). For MOA1, total precipitation depleted the isotopic signature of $\delta^{18}O$ by 0.01 ‰ for every mm of precipitation. The isotopic values of $\delta^{2}H$ had a similar depletion effect with a decrease of 0.06 ‰ in 2007 and 0.09 ‰ in 2008 for every mm of total precipitation.

VP had a positive correlation with the $\delta^{18}O$ and $\delta^{2}H$ values. For the 2007 data, VP increased $\delta^{18}O$ by 0.19 ‰ and $\delta^{2}H$ by 1.07 ‰ for each unit increase in hPa. VP was not included in the 2008 models as temperature was statistically more significant for that dataset. Nelson et al. (2021) discussed the effect of VP on the $\delta^{18}O$ and $\delta^{2}H$ values and attributed it to the effect of temperature on RH and its connection with the airmass reaching the temperatures of condensation and evaporation.

RH (%) had indicated a negative correlation with the $\delta^{18}O$ and $\delta^{2}H$ values. This is consistent with the fact that it has an impact on the sub-cloud evaporation which isotopically enriches precipitation (Dansgaard, 1964). For the 2007 dataset, relative humidity decreased $\delta^{18}O$ by 0.11 ‰ and $\delta^{2}H$ by 0.83 ‰ for every 1 % increase in RH. RH was not included in the models for 2008, as the models including temperature were statistically more significant.

Evapotranspiration had a positive correlation with the $\delta^{18}O$ and $\delta^{2}H$ values and enriched $\delta^{18}O$ by 0.03 ‰ and $\delta^{2}H$ by 0.32 ‰ for every 1 mm increase in evapotranspiration for the 2007 data. Dansgaard (1964) first discussed evaporation as an important factor and, along with temperature, noted it contributed to isotopic enrichment in precipitation. Evapotranspiration was not included in the models for 2008 as the models including temperature were statistically more significant.

Notably, evapotranspiration was only available from BoM as composite data, and thus, the monthly amounts were mean values. In Australia, evapotranspiration, along with the other climatic variables, has a strong spatial component, with values increasing as we move towards central Australia. This strong spatial component of the climatic covariates may also explain why other spatial covariates were less important in this study, for example temperature decreases with altitude and increases with latitude, and total precipitation also decreases moving inland, towards central Australia. This may indicate a longitudinal effect, that could be analysed in future studies.

[Include Figure 7 here]

4.4 Interpolation results

By using a leave-one-out analysis, it was found that both USK and USCo outperformed GLS, regardless of method of analysis. MOA1 proved to be the best method overall. This is probably because isotopic signatures have a very wide variation, therefore, the values seen in one year, do not necessarily help with the prediction of the values for the following year. The models from MOA1 (Equations (6), (7), (8) and (9) presented in Table 5) did not include any of the purely spatial covariates (distance, latitude, and altitude) suggesting, that the spatial aspect of the climatic covariates, are more important in the study of isotopic signature in southeastern Australia. Interestingly, previous studies (Erdélyi et al., 2023; Nelson et al., 2021) have also demonstrated the superior predictive power of climatic covariates.

[Include Table 5 here]

USK performed slightly better than USCo when analysing $\delta^{18}O$ ($RMSE$ 1.78 vs 1.82,

MAE 1.32 vs 1.35,$PRESS$ 1341 vs 1401) , and for $\delta^{2}H$, USK and USCo were both good models, ($RMSE$ 14.12 vs 14.19, MAE 10.78 vs 10.44, $PRESS$ 84751 vs 85533). Figure 4 illustrates how close the two methods were when predicting the observations. The small differences are practically negligible. When analysing the isoscapes created from USK, they appeared to have a smoother transition than those created using USCo, which sometimes displayed small points (bullseyes) around the observations (e.g., Figure 5, October 2008). This suggests that using USK was superior to USCo as kriging interpolation should smooth the transition between observations. This suggestion is supported by the goodness of fit values, which indicated USK is a slightly better fit. Notably, the slight improvement in MAS for $\delta^{2}H$, did not produce a smoother isoscape (see Table 6).

[Include Table 6 here]

The isoscapes illustrated how the $\delta^{18}O$ and $\delta^{2}H$ values changed in space, and the temporal correlation was clear. For example, the isoscapes in October, November and December were more similar than those in April, May, and June, for both 2007 and 2008. This confirms Tobler's (1970) first law of geography, i.e., close things are more similar than those farther apart. In June 2007, there were some very large precipitation events in Queensland, NSW and the Australian Capital Territory (ACT), with five east coast low pressure systems off the NSW coast. Brisbane and Charleville received 166 % and 131 % of their long-term average June precipitation respectively, and Canberra received 226 % of their long-term average June precipitation (BoM, 2023a). Many sites in NSW experienced their lowest recorded maximum temperature, and June 2007 was NSW's second coldest on record (BoM, 2023a). This cold and wet weather is a possible explanation for the very depleted values for $\delta^{18}O$ and $\delta^{2}H$ observed at Charleville and the ACT. The June 2007 isoscape illustrates how these events would have affected areas around these sites.

The kriging variance map was used to analyse how well the interpolation performed. Lower kriging variance values indicate a more accurate interpolation. Areas with a kriging variance of ≥ 4 (areas coloured red and yellow) were considered extrapolation, kriging variances < 4 were considered reliable (areas coloured blue and green, see Figure 5). As there were more collection sites around Sydney and towards the south of the study area, the interpolation was quite accurate, however, the north-western part of the study area, contains little data and thus, moves into extrapolation. Notably, the area south of Sydney is well documented however the isoscape would benefit from more data north of Sydney along the Great Dividing Range to Brisbane.

For this study USK outperformed USCo, it would be interesting to see if adding data from more years would improve the USCo interpolation. The algorithm for USCo took an exceptionally long time to execute (480 minutes vs 0.35 minutes for USK) and was a little unstable. Thus, interpolation had to be broken into monthly lots and concatenated for isoscape production. This was due to the inversion of large matrices at each prediction. Whilst R (R Core Team, 2022) can handle linear algebra equations with its built-in datatypes, an alternative approach could be to use the R packages **bigmemory** (Kane et al., 2013), and **biganalytics** (Emerson et al., 2020) which provide structures for working with matrices that are too large to fit into memory.

4.5 Other applications

Isoscapes can add value when calculating water budgets, and assessing flow pathways, transit times, and flux partitioning across the branches of the water cycle (Bowen & Good, 2015). They can identify the spatial patterns of shallow groundwaters (Bowen, 2010a); quantify components of the atmospheric water balance during large storms (Bennington & Farmer, 2014); and aid in the modelling of rainfall to runoff (McGuire et al., 2005) and surface water (Brooks et al., 2014; Jasechko

et al., 2013). Using the ST geostatistical approach and including the climate data from the same time-period as when $\delta^{18}O$and $\delta^{2}H$ values were observed, allows the water isotope hydrology researcher to reliably interpolate these values at unobserved points and create isoscapes that map not only the spatial aspect but also the temporal aspect of the data. For example, studies to reconstruct past hydroclimate have found that the $\delta^{2}H$ values of leaf wax are highly correlated with average annual precipitation $\delta^{2}H$ values globally, but not in the tropical Pacific (Ladd et al., 2021). Using the ST geostatistical approach to this research and incorporating the observed annual precipitation, rather than the annual average, may improve interpolation and identify correlations not yet found.

The data in this work was obtained from Australian sources. For other regions similar climatic data is available. For example, Bedaso et al. (2021) obtained climate data from the European Centre for Medium-Range Weather Forecasts. These data can easily be used with the existing code to create monthly isoscapes.

5 Conclusion

In conclusion, by using the ST geostatistical approach we were able to produce isoscapes on a month-to-month basis that produced a plausible interpolation between spatially and temporally distant observational data, which is superior to using that of the nearest observed site alone. The isoscapes produced clearly identified the ST correlation of the signatures with closer months and positions displaying more similarities to those farther apart. Both USK and USCo were better interpolation methods than simple linear GLS models and were able to predict $\delta^{2}H$ and $\delta^{18}O$ values more accurately at monthly intervals. For the east coast of Australia climatic covariates were more important for interpolation than other purely spatial variables. The best climatic covariates were total monthly precipitation, average monthly temperature, RH, VP, and evapotranspiration. Purely spatial covariates, like distance, altitude, and latitude, were not useful for prediction of monthly isotopic signature values. However, these spatial covariates may be useful for longer time series or at a coarser prediction interval (e.g., annual).

The isoscapes produced in this study, and future isoscapes produced using this method, will have application in provenance, forensics, and hydrology. Future study could include adding more data to this dataset to determine if USCo results improve and investigating the spatial nature of the climatic covariates. The methods and code developed and applied in this study can easily be used for the creation of new isoscapes using other datasets over different time periods and locations. Applications could include food provenance, forensics, and hydrology.

Data Availability Statement

The R code for universal spacetime cokriging is in supplementary materials Section E Part E.2. No original stable isotope data were generated for this study, and the reader is referred to the referenced sources. The gridded climate and spatial data are available in the supplementary materials (m.calc.csv).

Acknowledgments

Elevation data was calculated by Deanna Duffy, Spatial Data Analysis Network team, Charles Sturt University. This work was supported by an AINSE Honours Scholarship (2023), ANSTO and the Charles Sturt Honours and HECS exempt award. Thank you to Pauline Treble and Andy Baker for valuable comments on an early draft of this manuscript, and a special thanks to the AINSE WISE program and my long-term mentor Paticia Gadd for introducing me to ANSTO and my co-authors.

References

Allen, S. T., Kirchner, J. W., & Goldsmith, G. R. (2018). Predicting spatial patterns in precipitation isotope δ^2H and $\delta 1^8O$ seasonality using sinusoidal isoscapes. *Geophysical research letters, 45*(10), 4859-4868. https://doi.org/10.1029/2018gl077458

Araguás-Araguás, L., Froehlich, K., & Rozanski, K. (2000). Deuterium and oxygen-18 isotope composition of precipitation and atmospheric moisture. *Hydrological processes, 14*(8), 1341-1355. https://doi.org/10.1002/1099-1085(20000615)14:8<1341::AID-HYP983>3.0.CO;2-Z

Bedaso, Z., & Wu, S.-Y. (2021). Linking precipitation and groundwater isotopes in Ethiopia - Implications from local meteoric water lines and isoscapes. *Journal of Hydrology, 596*, 126074. https://doi.org/10.1016/j.jhydrol.2021.126074

Belkhiri, L., Tiri, A., & Mouni, L. (2020). Spatial distribution of the groundwater quality using kriging and co-kriging interpolations. *Groundwater for sustainable development, 11*, 100473. https://doi.org/10.1016/j.gsd.2020.100473

Bennington, J. B., & Farmer, E. C. (2014). *Learning from the impacts of superstorm Sandy*. Academic Press.

Birkel, C., Helliwell, R., Thornton, B., Gibbs, S., Cooper, P., Soulsby, C., Tetzlaff, D., Spezia, L., Esquivel-Hernández, G., Sánchez-Murillo, R., & Midwood, A. J. (2018). Characterization of surface water isotope spatial patterns of Scotland. *Journal of geochemical exploration, 194*, 71-80. https://doi.org/10.1016/j.gexplo.2018.07.011

Bivand, R. S., Pebesma, E. J., & Gomez-Rubio, V. (2013). *Applied spatial data analysis with R* (2nd ed.). Springer.

Bowen, G. J. (2010a). Isoscapes: spatial pattern in isotopic biogeochemistry. *Annual review of earth and planetary sciences, 38*, 161-187.

Bowen, G. J. (2010b). Statistical and geostatistical mapping of precipitation water isotope ratio. In J. B. West, G. J. Bowen, T. E. Dawson, & K. P. Tu (Eds.), *Isoscapes Understanding movement,*

pattern, and process on Earth through isotope mapping (1st ed. 2010. ed., pp. 425-429). Springer Netherlands. https://doi.org/10.1007/978-90-481-3354-3

Bowen, G. J., & Good, S. P. (2015). Incorporating water isoscapes in hydrological and water resource investigations. *Wiley Interdisciplinary Reviews: Water, 2*(2), 107-119.

Bowen, G. J., Guo, J. S., & Allen, S. T. (2022). A 3-D groundwater isoscape of the contiguous USA for forensic and water resource science. *PloS one, 17*(1), e0261651-e0261651. https://doi.org/10.1371/journal.pone.0261651

Bowen, G. J., & Revenaugh, J. (2003). Interpolating the isotopic composition of modern meteoric precipitation. *Water Resources Research, 39*(10), n/a-n/a. https://doi.org/10.1029/2003wr002086

Brooks, J. R., Gibson, J. J., Birks, S. J., Weber, M. H., Rodecap, K. D., & Stoddard, J. L. (2014). Stable isotope estimates of evaporation: inflow and water residence time for lakes across the United States as a tool for national lake water quality assessments. *Limnology and Oceanography, 59*(6), 2150-2165.

Bureau of Meteorology (BoM). (2023a). Australian Government. http://www.bom.gov.au/

Bureau of Meteorology (BoM). (2023b). *Maps of average conditions*. Australian Government. http://www.bom.gov.au/climate/averages/maps.shtml

Bureau of Meteorology (BoM). (2023c). *Australian Climate Influences*. Australian Government. http://www.bom.gov.au/watl/about-weather-and-climate/australian-climate-influences.shtml

Crawford, J., Hollins, S. E., Meredith, K. T., & Hughes, C. E. (2016). Precipitation stable isotope variability and subcloud evaporation processes in a semi-arid region. *Hydrological processes, 31*(1), 20-34. https://doi.org/10.1002/hyp.10885

Cressie, N. A. C. (1993). *Statistics for Spatial Data*. John Wiley & Sons, Incorporated.

689 Cressie, N. A. C., & Wikle, C. K. (2011). *Statistics for spatio-temporal data*. Wiley.

690

691 Crosbie, R., Morrow, D., Cresswell, R., Leaney, F., Lamontagne, S., & Lefournour, M. (2012). *New*
692 *insights to the chemical and isotopic composition of rainfall across Australia.* . CSIRO.
693 https://doi.org/https://doi.org/10.5072/83/5849a0d6c685a

694

695 Dansgaard, W. (1954). The O18-abundance in fresh water. *Geochimica et Cosmochimica Acta, 6*(5-6),
696 241-260.

697

698 Dansgaard, W. (1964). Stable isotopes in precipitation. *Tellus, 16*(4), 436-468.
699 https://doi.org/10.3402/tellusa.v16i4.8993

700

701 Delavau, C., Chun, K., Stadnyk, T., Birks, S. J., & Welker, J. M. (2015). N orth A merican precipitation
702 isotope (δ18O) zones revealed in time series modeling across C anada and northern U nited
703 S tates. *Water Resources Research*, *51*(2), 1284-1299.

704

705 Emerson, J. W., Kane, M. J., & Chandra , S. (2020). *biganalytics: Utilities for 'big.matrix' objects from*
706 *package 'bigmemory'.* In (Version 1.1.21) http://www.bigmemory.org

707

708 Erdélyi, D., Kern, Z., Nyitrai, T., & Hatvani, I. G. (2023). Predicting the spatial distribution of stable
709 isotopes in precipitation using a machine learning approach: a comparative assessment of
710 random forest variants. *GEM - International Journal on Geomathematics*, *14*(1), 14.
711 https://doi.org/10.1007/s13137-023-00224-x

712

713 ESRI. (2020). *ArcGIS Pro Version 2.7.0 [GIS software]* In Environmental Systems Research Institute.

714

715 Feng, X., Faiia, A. M., & Posmentier, E. S. (2009). Seasonality of isotopes in precipitation: A
716 global perspective. *Journal of Geophysical Research*, *114*(D8), D08116-n/a.
717 https://doi.org/10.1029/2008JD011279

718

719 Friedman, I. (1953). Deuterium content of natural waters and other substances. *Geochimica et*
720 *Cosmochimica Acta, 4*(1-2), 89-103. https://doi.org/10.1016/0016-7037(53)90066-0

721

Geoscience Australia. (2018). *Digital elevation model (DEM) of Australia derived from SRTM with 1 second grid - smoothed percentage slope WCS* [Elevation DEM].

González, J. A., Rodríguez-Cortés, F. J., Cronie, O., & Mateu, J. (2016). Spatio-temporal point process statistics: A review. *Spatial Statistics, 18*, 505-544. https://doi.org/https://doi.org/10.1016/j.spasta.2016.10.002

Graler, B., Pebesma, E. J., & Heuvelink, G. (2016). Spatio-temporal interpolation using gstat. *The R Journal, 8*(1).

Gribov, A., & Krivoruchko, K. (2020). Empirical Bayesian kriging implementation and usage. *Science of the Total Environment, 722*, 137290. https://doi.org/https://doi.org/10.1016/j.scitotenv.2020.137290

Hatvani, I. G., Erdélyi, D., Vreča, P., & Kern, Z. (2020). Analysis of the spatial distribution of stable oxygen and hydrogen isotopes in precipitation across the Iberian Peninsula. *Water (Basel), 12*(2), 481. https://doi.org/10.3390/w12020481

Hawkins, S., Northey, J., Schroder, I., Lem, A., Sedgmen, A., & Peljo, M. (2022). *Hydrochemistry database (HYDROCHEM)*.

Hollins, S. E., Hughes, C. E., Crawford, J., Cendón, D. I., & Meredith, K. T. (2018). Rainfall isotope variations over the Australian continent – Implications for hydrology and isoscape applications. *The Science of the total environment, 645*, 630-645. https://doi.org/10.1016/j.scitotenv.2018.07.082

Hu, D.-g., & Shu, H. (2019). Spatiotemporal interpolation of precipitation across Xinjiang, China using space-time CoKriging. *Journal of Central South University, 26*(3), 684-694. https://doi.org/10.1007/s11771-019-4039-1

Hughes, C. E., & Crawford, J. (2013). Spatial and temporal variation in precipitation isotopes in the Sydney Basin, Australia. *Journal of hydrology (Amsterdam), 489*, 42-55. https://doi.org/10.1016/j.jhydrol.2013.02.036

IAEA/WMO. (2022). *Global Network of Isotopes in Precipitation*.

Jasechko, S., Sharp, Z. D., Gibson, J. J., Birks, S. J., Yi, Y., & Fawcett, P. J. (2013). Terrestrial water fluxes dominated by transpiration. *Nature, 496*(7445), 347-350.

Jiao, Y., Liu, C., Gao, X., Xu, Q., Ding, Y., & Liu, Z. (2019). Impacts of moisture sources on the isotopic inverse altitude effect and amount of precipitation in the Hani Rice Terraces region of the Ailao Mountains. *The Science of the total environment, 687*, 470-478. https://doi.org/10.1016/j.scitotenv.2019.05.426

Kane, M. J., Emerson, J., & Weston, S. (2013). Scalable strategies for computing with massive data. *Journal of Statistical Software, 55*(14), 1-19.

Kaseke, K. F., Wang, L., Wanke, H., Turewicz, V., & Koeniger, P. (2016). An analysis of precipitation isotope distributions across Namibia using historical data. *PloS one, 11*(5), e0154598-e0154598. https://doi.org/10.1371/journal.pone.0154598

Kutner, M. H. (2005). *Applied linear statistical models* (B. Gordon, Ed. 5th ed.). McGraw-Hill Irwin.

Ladd, S. N., Maloney, A. E., Nelson, D. B., Prebble, M., Camperio, G., Sear, D. A., Hassall, J. D., Langdon, P. G., Sachs, J. P., & Dubois, N. (2021). Leaf Wax Hydrogen Isotopes as a Hydroclimate Proxy in the Tropical Pacific. *Journal of Geophysical Research: Biogeosciences, 126*(3), e2020JG005891. https://doi.org/https://doi.org/10.1029/2020JG005891

Lawrence, M. G. (2005). The Relationship between Relative Humidity and the Dewpoint Temperature in Moist Air: A Simple Conversion and Applications. *Bulletin of the American Meteorological Society, 86*(2), 225-233. https://doi.org/10.1175/BAMS-86-2-225

Liu, J., Fu, G., Song, X., Charles, S. P., Zhang, Y., Han, D., & Wang, S. (2010). Stable isotopic compositions in Australian precipitation. *Journal of Geophysical Research. Atmospheres, 115*(23). https://doi.org/10.1029/2010JD014403

Matiatos, I., & Wassenaar, L. I. (2019). Stable isotope patterns reveal widespread rainy-period-biased recharge in phreatic aquifers across Greece. *Journal of Hydrology, 568*, 1081-1092. https://doi.org/10.1016/j.jhydrol.2018.11.053

McGuire, K., McDonnell, J. J., Weiler, M., Kendall, C., McGlynn, B., Welker, J., & Seibert, J. (2005). The role of topography on catchment-scale water residence time. *Water Resources Research, 41*(5). W05002. https://10.1029/2004WR003657

Moreno, M., & Basille, M. (2018). *Drawing beautiful maps programmatically with R, sf and ggplot2*. R-spatial. Retrieved 6th September 2023 from https://r-spatial.org/r/2018/10/25/ggplot2-sf.html

Munksgaard, N. C., Kurita, N., Sánchez-Murillo, R., Ahmed, N., Araguas, L., Balachew, D. L., Bird, M. I., Chakraborty, S., Kien Chinh, N., Cobb, K. M., Ellis, S. A., Esquivel-Hernández, G., Ganyaglo, S. Y., Gao, J., Gastmans, D., Kaseke, K. F., Kebede, S., Morales, M. R., Mueller, M., . . . Zwart, C. (2019). Data Descriptor: Daily observations of stable isotope ratios of rainfall in the tropics. *Scientific reports, 9*(1), 14419. https://doi.org/10.1038/s41598-019-50973-9

Nelson, D. B., Basler, D., & Kahmen, A. (2021). Precipitation isotope time series predictions from machine learning applied in Europe. *Proceedings of the National Academy of Sciences - PNAS, 118*(26), 1. https://doi.org/10.1073/pnas.2024107118

Neto, J., Hidasi. (2014). How to create a grid and intersect it with a polygon/shape in R. *R Functions. A place to share and learn about R. Mainly destined to ecologists.* https://rfunctions.blogspot.com/2014/12/how-to-create-grid-and-intersect-it.html

Nlend, B., Huneau, F., Garel, E., Santoni, S., & Mattei, A. (2023). Precipitation isoscapes in areas with complex topography: Influence of large-scale atmospheric dynamics versus microclimatic phenomena. *Journal of Hydrology, 617*(128896). https://doi.org/10.1016/j.jhydrol.2022.128896

Pebesma, E. (2012). spacetime: Spatio-Temporal Data in R. *Journal of Statistical Software, 51*(7), 1 - 30. https://doi.org/10.18637/jss.v051.i07

Pebesma, E. (2023). *Package gstat. Spatial and spatio-temporal geostatistical modelling, prediction and simulation.* In https://github.com/r-spatial/gstat/

Pebesma, E. J. (2014). gstat user's manual. https://www.gstat.org/gstat.pdf

827 Pebesma, E. J. (2018). Simple features for R: Standardized support for spatial vector data. *The R*
828 *Journal, 10*(1), 439-446. https://doi.org/10.32614/RJ-2018-009

829

830 Pebesma, E. J., & Bivand, R. S. (2005a). Classes and methods for spatial data in R. *R News 5*(2). 9-13.

831

832 Pebesma, E. J., & Bivand, R. S. (2005b). *Classes and methods for spatial data: the sp package*.
833 Institute for Statistics and Mathematics of WU (Wirtschaftsuniversität Wien). https://cran.r-
834 project.org/

835

836 Pinheiro, J., & Bates, D. (2000). *Mixed-effects models in S and S-PLUS*. Springer.
837 https://doi.org/10.1007/b98882

838

839 Pinheiro, J., & Bates, D. (2023). *nlme: Linear and nonlinear mixed effects models.* In (Version R
840 package version 3.1-162) R Core Team. https://CRAN.R-project.org/package=nlme

841

842 R Core Team. (2022). *R: A language and environment for statistical computing.* In R Foundation for
843 Statistical Computing. https://www.R-project.org/

844

845 Ransley, T., Plazinska, A., Nation, E., Hostetler, S., Sundaram, B., & Brodie, R. S. (2015). *Cumulative*
846 *rainfall collector - A low-cost tool for assessing groundwater recharge. R*. ResearchGate.
847 Retrieved 29th September 2023 from
848 https://www.researchgate.net/publication/266413543_CUMULATIVE_RAINFALL_COLLECTO
849 R_-A_LOW-COST_TOOL_FOR_ASSESSING_GROUNDWATER_RECHARGE

850

851 Risi, C., Muller, C., & Blossey, P. (2020). What Controls the Water Vapor Isotopic Composition Near
852 the Surface of Tropical Oceans? Results From an Analytical Model Constrained by Large-Eddy
853 Simulations. *Journal of Advances in Modeling Earth Systems, 12*(8), e2020MS002106.
854 https://doi.org/https://doi.org/10.1029/2020MS002106

855

856 Rossiter, D. (2018). Co-kriging with the gstat package of the R environment for statistical computing.

857

858 Rostami, A. A., Karimi, V., Khatibi, R., & Pradhan, B. (2020). An investigation into seasonal variations
859 of groundwater nitrate by spatial modelling strategies at two levels by kriging and co-kriging
860 models. *Journal of environmental management, 270*, 110843-110843.
861 https://doi.org/10.1016/j.jenvman.2020.110843

Schabenberger, O., & Gotway, C. A. (2017). *Statistical Methods for Spatial Data Analysis* (First edition. ed.). Chapman and Hall/CRC.

South, A. (2017a). *rnaturalearth: World map data from Natural Earth. .* In (Version R package version 0.1.0) https://CRAN.R-project.org/package=rnaturalearth

South, A. (2017b). *_rnaturalearthdata, world vector map data from natural Earth used in 'rnaturalearth'_. R package version 0.1.0.* In https://CRAN.R-project.org/package=rnaturalearthdata

Subba Rao, T., & Terdik, G. (2017). A New Covariance Function and Spatio-Temporal Prediction (Kriging) for A Stationary Spatio-Temporal Random Process. *Journal of time series analysis, 38*(6), 936-959. https://doi.org/10.1111/jtsa.12245

Tadros, C. V., Markowska, M., Treble, P. C., Baker, A., Frisia, S., Adler, L. & Drysdale, R. N., (2022). Recharge variability in Australia's southeast alpine region derived from cave monitoring and modern stalagmite $\delta^{18}O$ records. *Quaternary Science Reviews*, *295*, 107742, https://doi.org/10.1016/j.quascirev.2022.107742

Terzer-Wassmuth, S., Wassenaar, L. I., Welker, J. M., & Araguás-Araguás, L. J. (2021). Improved high-resolution global and regionalized isoscapes of δ18O, δ2H and d-excess in precipitation. *Hydrological processes, 35*(6). https://doi.org/10.1002/hyp.14254

Terzer, S., Wassenaar, L. I., Araguás-Araguás, L. J., & Aggarwal, P. K. (2013). Global isoscapes for δ 18 O and δ 2 H in precipitation: improved prediction using regionalized climatic regression models. *Hydrology and earth system sciences, 17*(11), 4713-4728. https://doi.org/10.5194/hess-17-4713-2013

Thompson, J. A., Roecker, S., Grunwald, S., & Owens, P. R. (2012). Chapter 21 - Digital Soil Mapping: Interactions with and Applications for Hydropedology. In H. Lin (Ed.), *Hydropedology* (pp. 665-709). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-386941-8.00021-6

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Econ.Geogr, 46*, 234-240. https://doi.org/https://doi.org/10.2307/143141

Wackernagel, H. (2010). *Multivariate geostatistics : An introduction with applications* (3rd, completely rev. ed.). Springer.

Wang, S., Lei, S., Zhang, M., Hughes, C., Crawford, J., Liu, Z., & Qu, D. (2022). Spatial and seasonal isotope variability in precipitation across China: monthly isoscapes based on regionalized fuzzy clustering. *Journal of climate, 35*(11), 3411-3425. https://doi.org/10.1175/JCLI-D-21-0451.1

Wassenaar, L. I., Van Wilgenburg, S. L., Larson, K., & Hobson, K. A. (2009). A groundwater isoscape (δD, δ 18O) for Mexico. *Journal of geochemical exploration, 102*(3), 123-136. https://doi.org/10.1016/j.gexplo.2009.01.001

Wickham, H., & Grolemund, G. (2017). *R for data science : import, tidy, transform, visualize, and model data* (First edition. ed.). O'Reilly.

Wikle, C. K., Zammit-Mangion, A., & Cressie, N. (2019). *Spatio-temporal statistics with R*. Chapman & Hall/CRC.

Willmott, C., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research, 30*, 79–82. https://doi.org/10.3354/cr030079

Xu, Q., Jiao, Y., Liu, C., Liu, Z., Ding, Y., Zhang, H., Tao, Y., & Zhang, Z. (2021). The spatial patterns and impact factors of stable oxygen and hydrogen isoscapes in pond water: A case study on the water-source forests of the Hani terraced fields in Yunnan, China. *Journal of Hydrology 603*, 127097. https://doi.org/10.1016/j.jhydrol.2021.127097

Tables

Table 1 Summary of climatic variables included in rainfall isotope analysis models (ET -> Evapotranspiration, Temp. -> Temperature). ML refers to machine learning, UK refers to studies that considered UK but rejected its use for various reasons.*

Study	Dewpoint	Latent Heat Flux	Long or Short-wave radiation	Total Rainfall	Relative humidity	Vapour pressure	Wind speed	ET	Temp.	Interpolation Method
Bedaso & Wu, 2021				✓					✓	UC
Bowen and Good (2015)				✓				✓	✓	UK
Bowen et al., 2022				✓						3-D kriging
Delavau et al. 2015				✓	✓			✓		Regression
Dudley et al. 2024)			✓	✓	✓	✓	✓	✓	✓	UK
Erdélyi et al., 2023				✓		✓			✓	ML
Hatvani et al., 2020				✓					✓	UK*
Hollins et al., 2018				✓	✓	✓			✓	UK*
Kaseke et al., 2016				✓	✓			✓	✓	UC
Krajcar Bronić and Barešić (2021)				✓				✓	✓	UK
Nelson et al., 2021			✓	✓	✓	✓	✓	✓	✓	ML
Terzer-Wassmuth et al., 2021		✓	✓	✓	✓	✓	✓		✓	UK*

Xu et al. (2021)				⚑				⚑	⚑	UK
This study	⚑			⚑	⚑	⚑		⚑		USK, USCo

929

930 *Table 2 List of sites in study area.*

Site	Longitude	Latitude	Altitude (MASL)	Distance (km)	Observations	Source	Citation
Albury (Bungowannah)	146.76	-36.02	240	240	6	Geoscience Australia	Hawkins et. al., (2022); Ransley et al. (2015)
Barakula	150.50	-26.43	349	252	13	Geoscience Australia	Hawkins et. al., (2022); Ransley et al. (2015)
Big Hill	150.00	-34.57	652	75	19	ANSTO	Hughes & Crawford (2013)
Braidwood	149.80	-35.45	49	49	15	Geoscience Australia	Hawkins et. al., (2022); Ransley et al. (2015)
Brisbane (Aerodrome)	153.08	-27.43	4	4	24	CSIRO, Land and Water - Adelaide	Crosbie et al. (2012)
Charleville (Aerodrome)	146.26	-26.41	302	554	21	CSIRO, Land and Water - Adelaide	Crosbie et al. (2012)
Cobar	145.83	-31.48	260	554	22	CSIRO, Land and Water - Adelaide	Crosbie et al. (2012)
Condobolin	147.15	-33.08	375	375	4	Geoscience Australia	Hawkins et. al., (2022); Ransley et al. (2015)
Cowra	148.70	-33.80	212	212	7	Geoscience Australia	Hawkins et. al., (2022); Ransley et al. (2015)
Deniliquin	144.95	-35.53	260	260	5	Geoscience Australia	Hawkins et. al., (2022); Ransley et al. (2015)

Dubbo	148.61	-32.24	286	286	5	Geoscience Australia	Hawkins et. al., (2022); Ransley et al. (2015)
Evatt	149.08	-35.20	117	117	12	Geoscience Australia	Hawkins et. al., (2022); Ransley et al. (2015)
Gilgandra	148.66	-31.71	318	318	5	Geoscience Australia	Hawkins et. al., (2022); Ransley et al. (2015)
Googong (Queanbeyan Bowling Club)	149.23	-35.36	97	97	12	Geoscience Australia	Hawkins et. al., (2022); Ransley et al. (2015)
Goondiwindi	150.36	-28.60	302	302	6	Geoscience Australia	Hawkins et. al., (2022); Ransley et al. (2015)
Horsham	142.30	-36.67	1	182	12	Geoscience Australia	Hawkins et. al., (2022); Ransley et al. (2015)
Lithgow	150.17	-33.46	966	109	22	ANSTO	Hughes & Crawford (2013)
Lucas Heights	150.98	-34.05	152	15	23	ANSTO	Hughes & Crawford (2013)
Macquarie Marshes	147.49	-30.89	153	528	11	ANSTO	Crawford et al. (2016)
Melbourne (Airport)	144.97	-37.82	28	72	24	CSIRO, Land and Water - Adelaide	Crosbie et al. (2012)
Mildura (Airport)	142.09	-34.24	50	307	23	CSIRO, Land and Water - Adelaide	Crosbie et al. (2012)
Moree (Pallamallawa)	150.14	-29.47	303	303	8	Geoscience Australia	Hawkins et. al., (2022); Ransley et al. (2015)
Mt Werong	149.91	-34.08	1178	97	24	ANSTO	Hughes & Crawford (2013)
Narrabri	149.78	-30.32	310	310	4	Geoscience Australia	Hawkins et. al., (2022); Ransley et al. (2015)

Nyngan	147.20	-31.55	439	439	5	Geoscience Australia	Hawkins et. al., (2022); Ransley et al. (2015)
Sydney (Airport)	151.17	-33.94	6	4	24	CSIRO, Land and Water - Adelaide	Crosbie et al. (2012)
Tarcutta	147.74	-35.28	227	227	7	Geoscience Australia	Hawkins et. al., (2022); Ransley et al. (2015)
Tidbinbilla	148.94	-35.44	118	118	14	Geoscience Australia	Hawkins et. al., (2022); Ransley et al. (2015)
Toowoomba	151.88	-27.52	117	117	6	Geoscience Australia	Hawkins et. al., (2022); Ransley et al. (2015)
Wagga Wagga	147.46	-35.16	212	255	24	CSIRO, Land and Water - Adelaide	Crosbie et al. (2012)
Warialda	150.58	-29.54	258	258	6	Geoscience Australia	Hawkins et. al., (2022); Ransley et al. (2015)
Yarrangobilly	148.50	-35.73	990	143	12	ANSTO	Tadros et al. (2022)

931

932

933 *Table 3 Summary of covariates used in study. Calculated values are denoted with an asterisk.*

Data	Variable	Description	Unit
Dewpoint*	DP	Average monthly	°C
Evapotranspiration	evap	Average areal actual (composite)	mm month^{-1}
Maximum Temperature	temp	Average monthly (2007-2008)	°C
Precipitation	rain	Total monthly (2007-2008)	mm month^{-1}
Relative Humidity *	RH	Average monthly	%
Vapour Pressure	VP	Average monthly (2007-2008)	hPa
Altitude	alt	Meters above sea level	MASL
Distance*	dist	Distance from the coast	km
Slope	slope	Change in slope	%
Latitude	lat	Decimal degrees of Latitude	°

934

935 *Table 4 Pearson's correlation matrix for all covariates.*

	$\delta^{18}O$	$\delta^{2}H$	rain	evap	VP	RH	dist	slope	lat	Alt	DP	temp
$\delta^{18}O$	1.000											
$\delta^{2}H$	0.979	1.000										
rain	-0.224	-0.192	1.000									
evap	0.272	0.292	0.310	1.000								

VP	0.206	0.193	0.412	0.667	1.000							
RH	-0.445	-0.406	0.412	0.052	0.123	1.000						
dist	0.215	0.149	-0.257	-0.281	-0.130	-0.620	1.000					
slope	-0.206	-0.169	0.086	0.026	-0.221	0.203	-0.158	1.000				
lat	0.206	0.179	0.072	0.118	0.331	-0.196	0.408	-0.156	1.000			
alt	-0.217	-0.180	0.139	0.017	-0.227	0.224	-0.023	0.704	-0.027	1.000		
DP	0.212	0.199	0.400	0.643	0.989	0.085	-0.124	-0.231	0.296	-0.240	1.000	
temp	0.496	0.454	0.007	0.417	0.630	-0.676	0.393	-0.319	0.375	-0.328	0.664	1.000

936

937 *Table 5 Models derived from linear analysis. Where evap is evapotranspiration, rain is total monthly*
938 *precipitation, and $\boldsymbol{\beta}$ is the vector of regression coefficients.*

Year	Models	Equation#
2007	$\delta^{18}\hat{O}_{2007} = \beta_1 + \beta_2 RH + \beta_3 \text{evap} + \beta_4 rain + \beta_5 VP$	(6)
	$\widehat{\delta^2 H_{2007}} = \beta_1 + \beta_2 RH + \beta_3 \text{evap} + \beta_4 rain + \beta_5 VP$	(7)
2008	$\widehat{\delta^{18} O_{2008}} = \beta_1 + \beta_2 temp + \beta_3 rain$	(8)
	$\widehat{\delta^2 H_{2008}} = \beta_1 + \beta_2 temp + \beta_3 rain$	(9)
2007-2008	$\widehat{\delta^{18} O_{2007,2008}} = \beta_1 + \beta_2 RH + \beta_3 evap + \beta_4 rain + \beta_5 alt + \beta_6 lat + \beta_7 lat^2$	(10)
	$\widehat{\delta^2 H_{2007,2008}} = \beta_1 + \beta_2 RH + \beta_3 evap + \beta_4 rain + \beta_5 alt + \beta_6 lat + \beta_7 lat^2$	(11)

939

940 *Table 6 Statistics calculated from three methods. AIC and BIC were not calculated for Method 1 as it*
941 *implemented separate models for each year.*

Response	Model	Method	RMSE	MAS	PRESS	AIC	BIC
$\delta^{18}O$	GLS	MOA1	2.319859	1.753587	2287		
		MOA2	2.70544	2.039856	3111	858	882
		MOA3	2.667047	1.990873	3023	834	858
	USCo	MOA1	1.81586	1.345082	1401		
		MOA2	2.272156	1.636905	2194	710	734
		MOA3	2.285515	1.647153	2220	707	731
	USK	MOA1	1.776263	1.321963	1341		
		MOA2	1.988303	1.497501	1680	596	621
		MOA3	1.954975	1.51377	1578	564	590
$\delta^{2}H$	GLS	MOA1	18.31306	13.91761	142532		
		MOA2	21.41042	16.26885	194823	2616	2641
		MOA3	21.04792	15.75802	188281	2540	2565
	USCo	MOA1	14.18643	10.44189	85533		
		MOA2	17.52852	12.46408	130581	2446	2471
		MOA3	17.74843	12.59087	133878	2400	2424
	USK	MOA1	14.12139	10.78883	84751		
		MOA2	18.54572	13.59346	146176	2494	2519
		MOA3	15.5585	11.87783	99974	2279	2303

942

943

944

945 **Figure Captions**

946 Figure 1 Study area and Seasonal rainfall zones of Australia. Red triangles represent collection sites,
947 grey area (outlined with the red boarders), is the Murray Darlin Basin. Map produced with ArcGis Pro
948 (ESRI, 2020); base maps: (ESRI, 2009;2015). Rainfall map based on median annual rainfall 1900-1999
949 (BoM, 2023, Creative Commons 4.0)

950 Figure 2 Scatter plot of $\delta^{18}O$ over the period 2007-2008, red line indicates line of best fit, for this
951 dataset; the grey area includes 95% confidence interval.

952 Figure 3 Comparison of residuals and lagged residuals (Residuals -1) for all models.

953 Figure 4 Comparison of USK (top) and USCo (bottom) results derived from MOA1. The red lines (top)
954 indicate USK results, the blue lines (bottom) indicate USCo results, and the dashed black lines
955 indicate the GLS results.

956 Figure 5 Isoscapes for $\delta^{18}O$, (left) created using USK results, (centre) created using USCo results,
957 (right) kriging variance map.

958 Figure 6 Spatial covariates vs $\delta^{2}H$. (A) Altitude, (B) altitude with outlier removed, (C) slope, and (D)
959 slope with outliers removed and (E) Latitude. The red lines indicate the linear relationship between
960 the observed values of $\delta^{2}H$ and the covariate. The blue line in graph (E) indicates the parabolic
961 relationship from MOA2 $\widehat{\delta^{2}H} = mean(lat) + \beta_1 |lat| - \beta_2 lat^2$

962 Figure 7 Scatterplots comparing climatic covariates and $\delta^{2}H$ – the red line indicates the linear
963 relationship.

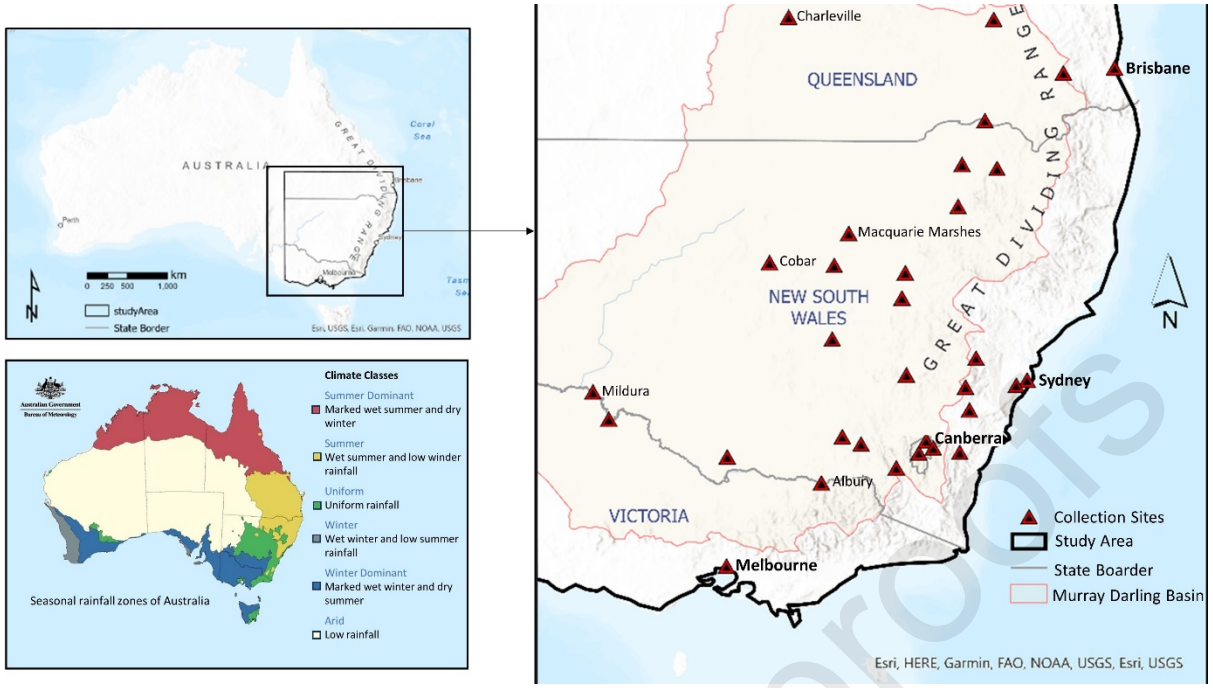
964

965

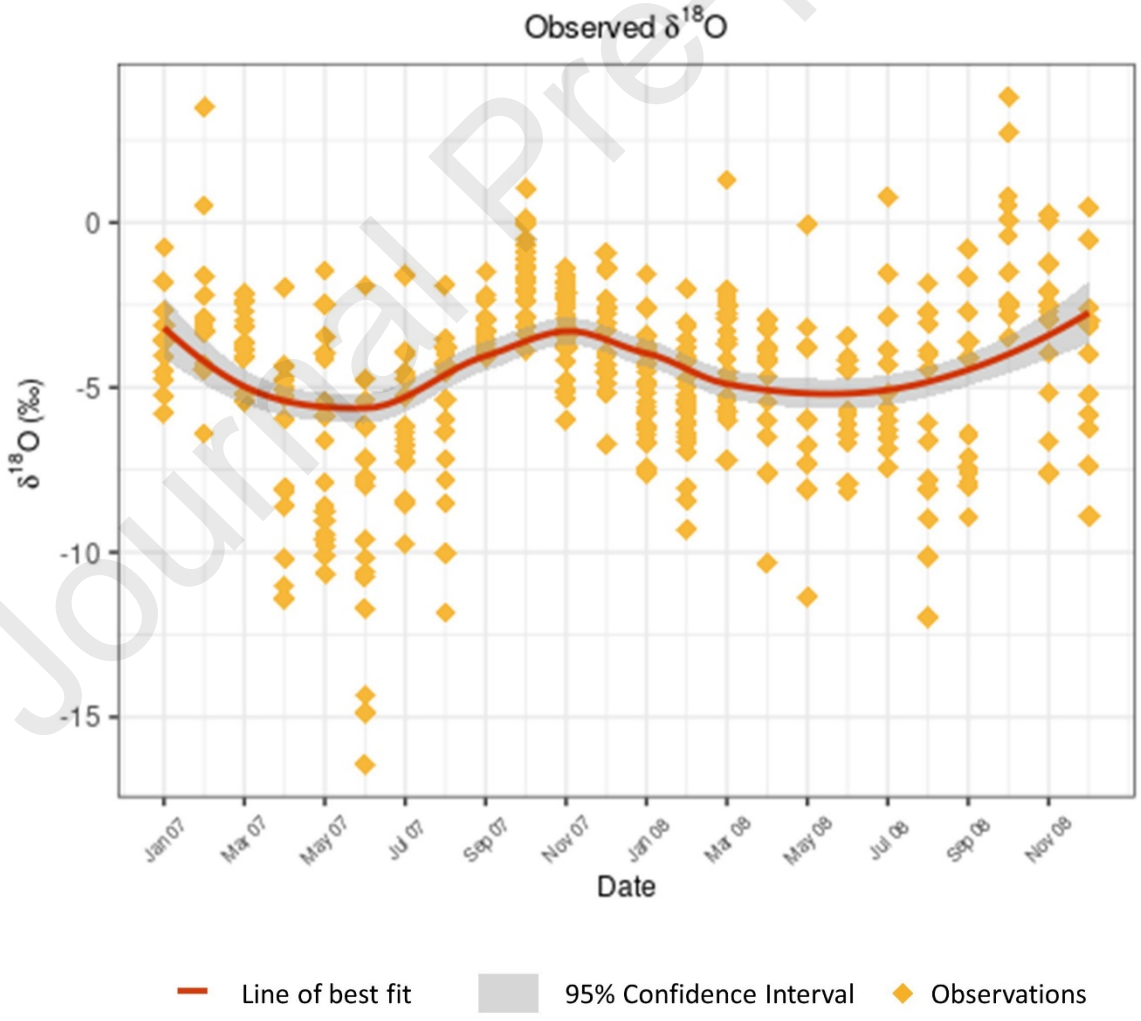
966

967

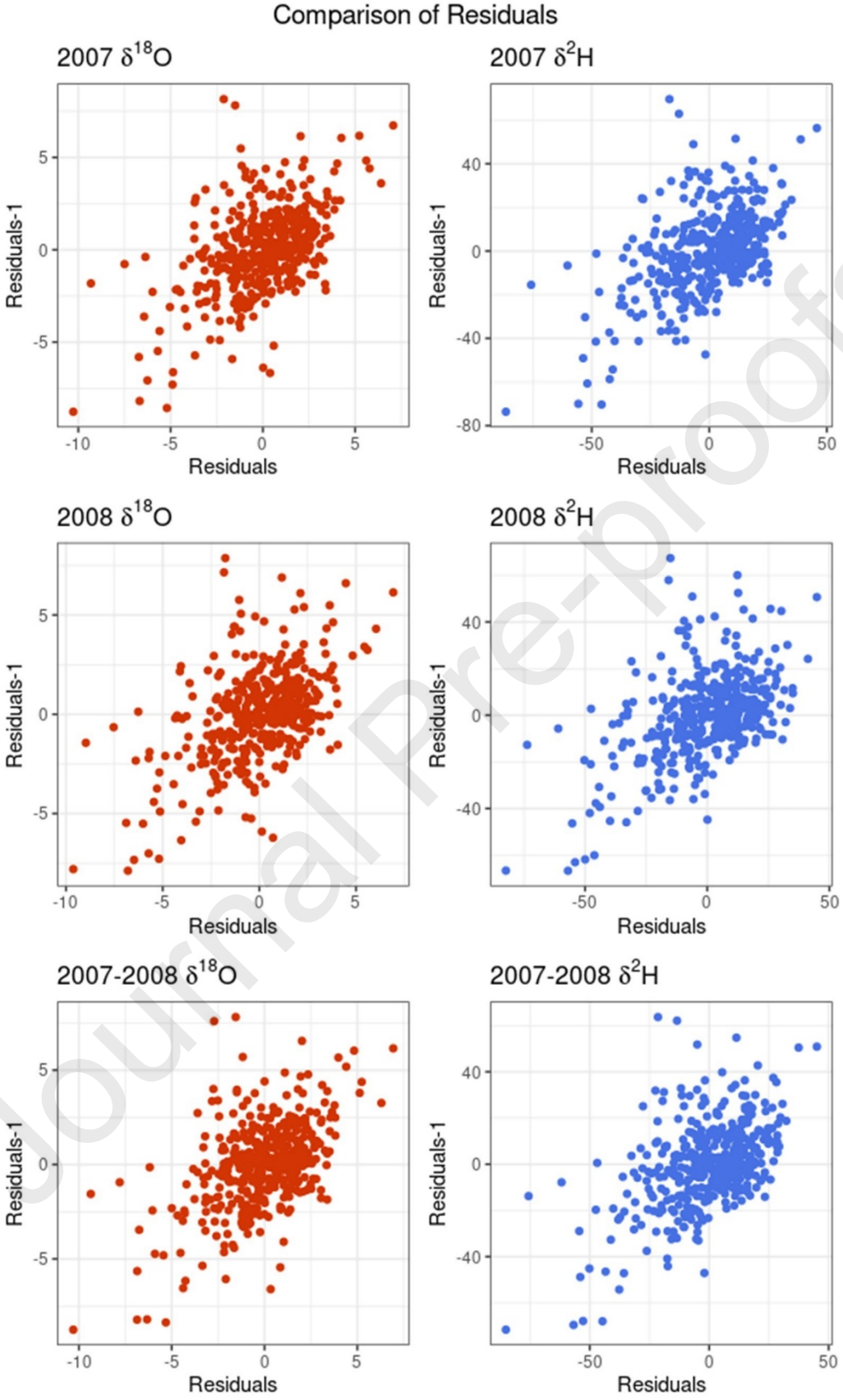
968


AUSTRALIA
Coral Sea
Perth
Brisbane
Sydney
km
0 250 500 1,000
studyArea
State Border
Esri, USGS, Esri, Garmin, FAO, NOAA, USGS
Climate Classes
Summer Dominant
Marked wet summer and dry winter
Summer
Wet summer and low winder rainfall
Uniform
Uniform rainfall
Winter
Wet winter and low summer rainfall
Winter Dominant
Marked wet winter and dry summer
Arid
Low rainfall
Seasonal rainfall zones of Australia
Charleville
QUEENSLAND
Brisbane
GREAT DIVIDING RANGE
Macquarie Marshes
Cobar
NEW SOUTH WALES
Mildura
Sydney
Canberra
Albury
VICTORIA
Melbourne
N
Collection Sites
Study Area
State Boarder
Murray Darling Basin
Esri, HERE, Garmin, FAO, NOAA, USGS, Esri, USGS

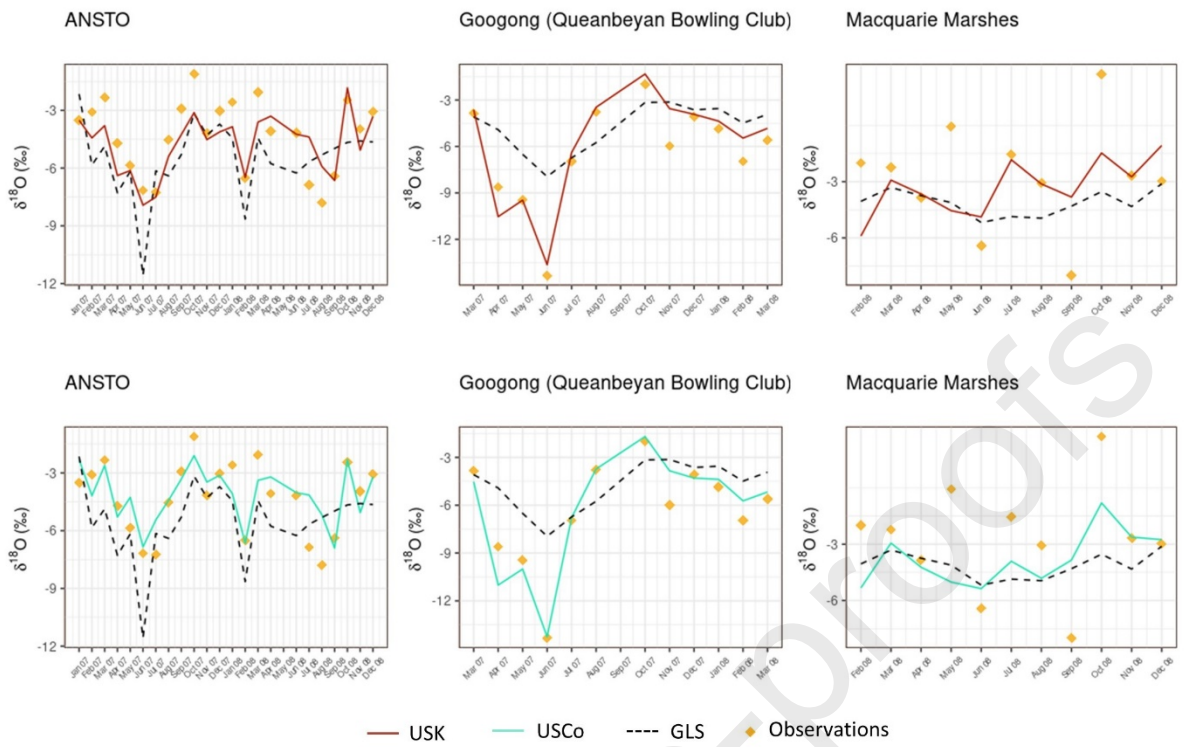
969


Observed δ18O
δ18O (‰)
0
-5
-10
-15
Jan 07
Mar 07
May 07
Jul 07
Sep 07
Nov 07
Jan 08
Mar 08
May 08
Jul 08
Sep 08
Nov 08
Date
Line of best fit
95% Confidence Interval
Observations

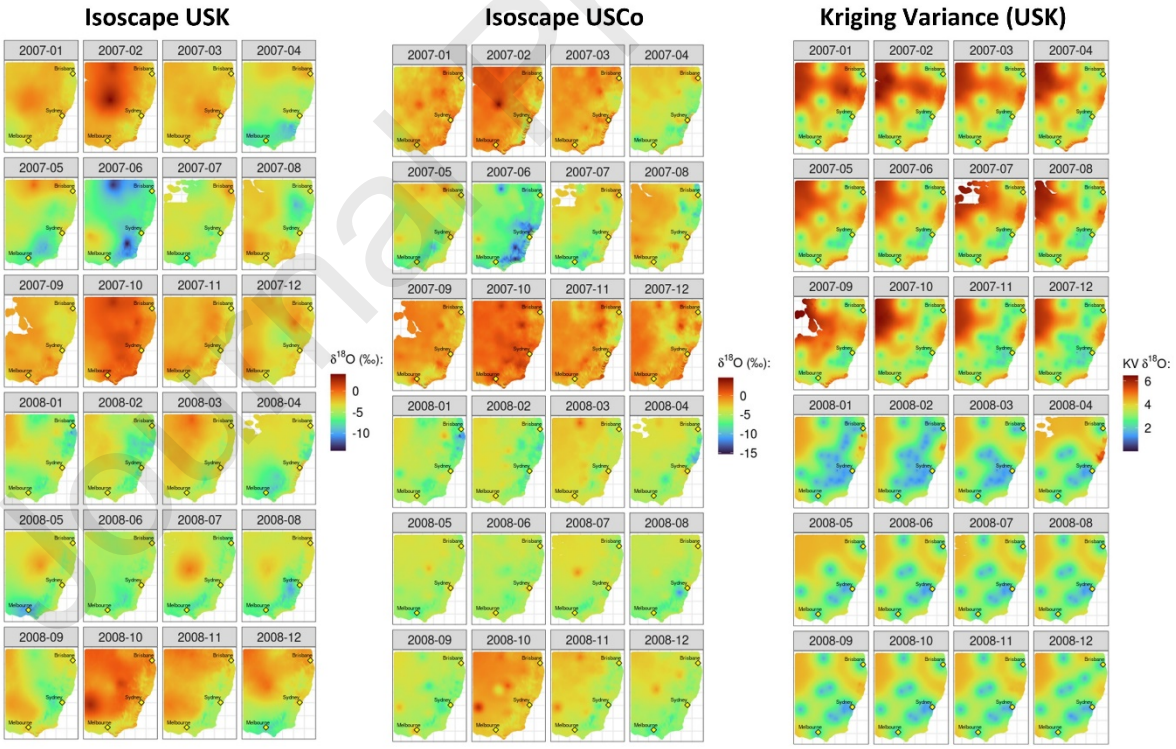
970


Comparison of Residuals
2007 δ18O
2007 δ2H
2008 δ18O
2008 δ2H
2007-2008 δ18O
2007-2008 δ2H
Residuals-1
Residuals

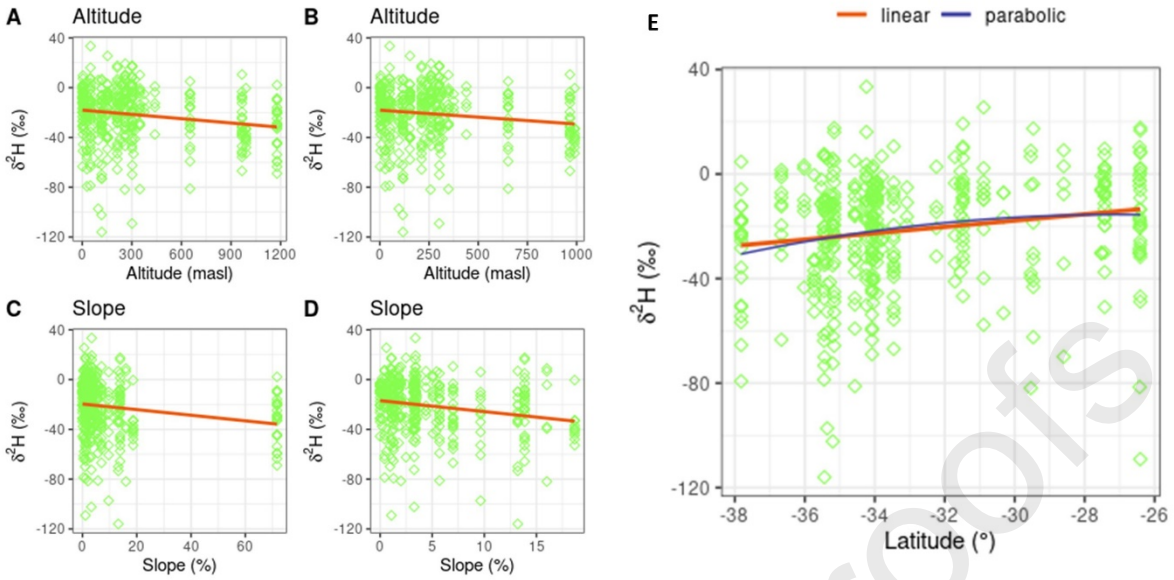
971


ANSTO
Googong (Queanbeyan Bowling Club)
Macquarie Marshes
δ18O (‰)
ANSTO
Googong (Queanbeyan Bowling Club)
Macquarie Marshes
δ18O (‰)
USK
USCo
GLS
Observations

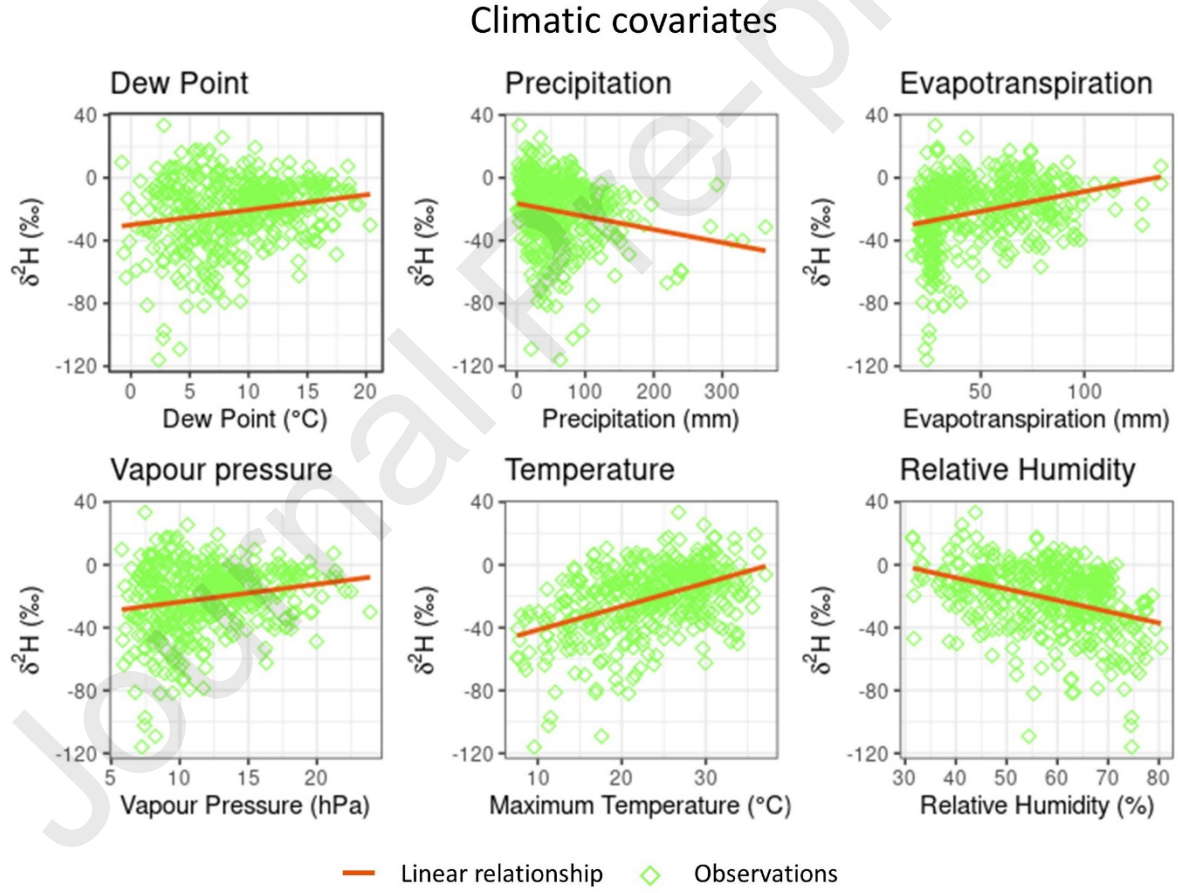
972


Isoscape USK
Isoscape USCo
Kriging Variance (USK)
δ18O (‰):
δ18O (‰):
KV δ18O:

973



974



975

Using spacetime geostatistical analysis to improve precipitation isoscape interpolation in Australia

Context

Precipitation isotope signatures often display temporal correlation – leading to a need for spatio-temporal interpolation methods.

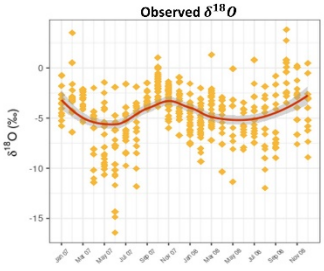


Fig.1 $\delta^{18}O$ observations 2007-2008. Red line indicates line of best fit, grey area includes 95% confidence interval

Methods

Universal spatio-temporal kriging (USK) and co-kriging (USCo) were used to interpolate values in a leave-one-out analysis and compared with generalised lease squares (GLS) interpolation.

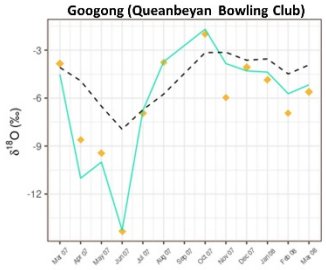


Fig.2 Comparison using leave-one-out analysis. Blue line represents results from USCo, black dashed line GLS and orange diamonds observed values.

Outcomes

USK and USCo algorithm greatly improve interpolation. Monthly isoscapes for south-eastern Australia created. R code for the USCo algorithm has been made available

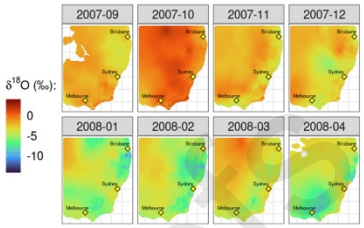


Fig.3 Monthly isoscapes created from USK for $\delta^{18}O$, blank areas refer to months with zero rainfall.

Conclusion

Using month-to-month local climate covariates produce plausible isoscapes that can be used in provenance, forensics, and hydrology applications. Universal spatio-temporal kriging (USK) and co-kriging (USCo) provided superior interpolation and were able to better predict the signatures of $\delta^{2}H$ and $\delta^{18}O$, at monthly intervals, when compared to other models such as generalised least squares (GLS). The spatio-temporal correlation was clearly identified. This method of interpolation can be easily adapted and used with different datasets.

Acknowledgments

This study was supported by the AINSE Honours Scholarship, ANSTO Australia and CSU Honours Scholarship HECS exempt award.

Reference

Duff, C. M., Crawford, J., Ip, R. H. L., Li, Z., Hughes, C. E., & Tadros, C. V. (2023). *Using spacetime geostatistical analysis to improve precipitation isoscape interpolation in Australia*. Charles Sturt University.

Highlights

- Algorithm for universal spacetime cokriging was developed for R.
- Universal spacetime kriging and cokriging improved interpolation.
- Climatic covariates more influential than purely spatial covariates.