

Comparison of ERA-40, ERA-Interim and NCEP/NCAR reanalysis data with observed surface air temperatures over Ireland

P. A. Mooney,^{a*} F. J. Mulligan,^b and R. Fealy^a

^a ICARUS, NUI Maynooth, Kildare, Ireland

^b Department of Experimental Physics, NUI Maynooth, Kildare, Ireland

ABSTRACT: Surface air temperatures modelled by ERA-40, ERA-Interim and (NCEP)/(NCAR) reanalysis (NNRP-1) have been compared with observations at 11 synoptic stations in Ireland over the period 1989–2001. The three reanalysis datasets show good agreement with the observed data and with each other. Slopes of the least-squares line to scatter plots of reanalysis data *versus* observational data show small differences between the three reanalyses, with ERA-40, ERA-Interim and NNRP-1 slopes ranging between $(0.79–1.06) \pm 0.01$, $(0.83–1.01) \pm 0.01$ and $(0.76–0.98) \pm 0.01$, respectively. Summary statistics and the monthly mean temperatures over the 1989–2001 period showed that the reanalyses were significantly warmer in winter than the observations, which resulted in best fit lines with slopes consistently less than unity. ERA-Interim was slightly better than both ERA-40 and NNRP-1 at modelling winter temperatures and it had higher correlation coefficients with the observations. All three reanalyses use different grid sizes and types. Subsequent regridding of the ERA-Interim and NNRP-1 data to the ERA-40 grid showed that the grid difference had no significant influence on the results. Comparison of ERA-Interim and NNRP-1 data with the air temperatures at four marine buoys around the Irish coast for the period 2001–2005 showed that the reanalyses modelled colder winter temperatures than the observations; resulting in best fit lines with slopes consistently greater than unity. The slopes for NNRP-1 and ERA-Interim at the marine buoys, respectively, averaged 1.09 ± 0.04 and 1.10 ± 0.05 while the slopes at the four land stations over the same period averaged 0.87 ± 0.02 and 0.89 ± 0.02 , respectively. We believe that this pattern results from the difference in the treatment of land and sea surfaces in the reanalysis datasets. Copyright © 2010 Royal Meteorological Society

KEY WORDS reanalysis data; observed data; surface air temperatures; data assimilation; marine buoys; Ireland; regridding

Received 21 July 2009; Revised 18 December 2009; Accepted 6 January 2010

1. Introduction

Over the past decade, reanalysis data has found widespread application in many areas of research ranging from studies of climatic trends (Poveda *et al.*, 2006; Ciccarelli *et al.*, 2008; Stammerjohn *et al.*, 2008) and climate modelling (Wang *et al.*, 2006; Fealy and Sweeney, 2007; Ciccarelli *et al.*, 2008) to estimation of renewable energy resources (Khan and Iqbal, 2004; Henfridsson *et al.*, 2007). It is advantageous to use reanalysis data in certain research areas where observational data are sparse or when knowledge of the state of the atmosphere on a uniform grid is required.

Reanalysis products are generated by the assimilation of observational data over a given period of time. Data assimilation entails the incorporation of observations into a background field to produce an initial condition. Techniques for assimilating these data have advanced greatly over the past few decades and two of the most popular methods are 3D- and 4D-variational analyses (Kalnay, 2007).

Although assimilation techniques continue to improve over time, a given reanalysis dataset uses the same assimilation method for the entire period of coverage. This avoids the possibility of variations in the reanalysis products because of the use of different assimilation systems (the possibility of variations because of any changes in the system of observations would of course remain). Reanalysis data are typically separated into three major phases corresponding to the evolution of the observing system (Kistler *et al.*, 2001): the ‘early’ period from 1940 to 1957 when the first upper air observations were established; the era of the ‘modern rawinsonde network’ from 1958 to 1978; and the ‘modern satellite’ era from 1979 to present.

Some of the most well-known reanalysis datasets are National Centers for Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR) (NNRP-1), NCEP/DOE (NNRP-2), ERA-15, ERA-40 and ERA-Interim. The variable examined in this study is the 2 m surface air temperature. Because the observational data used by the three reanalysis datasets are largely identical, it has frequently been assumed that any one of these datasets are equally valid for subsequent

*Correspondence to: P. A. Mooney, ICARUS, NUI Maynooth, Kildare, Ireland. E-mail: priscilla.a.mooney@nuim.ie

climate analysis or climate change prediction. However, previous studies have shown subtle differences between the NNRP-1 and ERA-40 reanalysis datasets (Escoffier and Provost, 1998; Li *et al.*, 2004; Ruti *et al.*, 2008). These small differences can have a significant impact on climate studies, e.g. solar signals in the atmosphere (Gleiser *et al.*, 2005) or drought variability (Bordi *et al.*, 2006).

Simmons *et al.* (2004) compared global surface air temperature anomalies from CRUTEM2v (Climatic Research Unit Temperature version 2 variance corrected; Jones and Moberg, 2003) with ERA-40 and NNRP-1. They found that ERA-40 showed better agreement with CRUTEM2v than NNRP-1 over the time periods 1958–2001 and 1979–2001. In addition, they reported closer agreement between ERA-40 and CRUTEM2v over the period 1979–2001 compared with the period 1958–1979. This was attributed to the greater observational coverage after 1979 which is linked to the inclusion of satellite data.

A study on the surface air temperature trends in both NNRP-1 and station data in the continental United States was conducted by Kalnay and Cai (2003). Their study showed that the NNRP-1 temperatures became cooler with time compared with the temperatures observed at the stations. The authors suggested that this difference was because of land use change, such as urbanisation, agriculture and irrigation.

The study reported here compares the 2 m air temperature in the ERA-40, ERA-Interim and NNRP-1 datasets with observational data over Ireland for the period 1989–2001. This variable is examined because of its importance to ecosystems, agriculture and tourism industries, and there is a substantial quantity of observational data available. The period 1989–2001 was chosen because all four datasets have data during this period.

ERA-Interim and NNRP-1 data are also compared with observational data from four marine buoys situated around the Irish coast for the period 2001–2005. The period of study for these is rather limited because the marine buoys were only deployed after 2001. ERA-40 reanalysis data were not included in this part of the investigation because its dataset terminates in 2002.

2. Surface temperature datasets

2.1. ERA-40 data

ERA-40 data (Uppala *et al.*, 2005) was produced by the European Center for Medium range Weather Forecasting (ECMWF) and it covers the period from 1958 to 2001. The assimilation method used in the ERA-40 project was 3D-variational analysis. It used a spectral grid with triangular truncation of 159 waves (corresponds approximately to 125 km) and a hybrid vertical coordinate system with 60 levels. The ECMWF global model was used for the background forecast. The ERA-40 data used in this study were obtained from the ECMWF data

server on a fixed grid of 2.5° resolution (corresponds to approximately 250 km).

2.2. ERA-Interim

ERA-Interim (Simmons *et al.*, 2006), also produced by ECMWF, uses 4D-variational analysis on a spectral grid with triangular truncation of 255 waves (corresponds to approximately 80 km) and a hybrid vertical coordinate system with 60 levels. The ECMWF global model is used for the forward integration in the 4D-variational analysis and the temporal length of the variational window is 12 h. This reanalysis covers the period from 1989 to present day. The ERA-Interim data used in this study was obtained from the ECMWF data server on a fixed grid of 1.5°.

2.3. NCEP/NCAR Reanalysis data

NCEP/NCAR Reanalysis Products-1 (NNRP-1) data was produced by the National Centers for Environmental prediction (NCEP) in collaboration with the National Centre for Atmospheric Research (NCAR) and it covers the period from 1948 to present day. The data assimilation system uses a 3D-variational analysis scheme, with 28 sigma levels in the vertical and a triangular truncation of 62 waves which corresponds to a horizontal resolution of approximately 200 km. A forecast from the NCEP global spectral model was used as the first-guess fields for this reanalysis. The assimilation system used in NCEP/NCAR reanalysis is described in more detail in Kalnay *et al.* (1996).

2.4. Station data

The station data used in this study was obtained from Met Éireann, the Irish National Meteorological Service, for 11 synoptic weather stations operated by Met Éireann for the period 1961–2005. The synoptic stations, which are geographically dispersed around Ireland, represent the conditions for a mixture of both coastal and inland locations. The synoptic network is manned by experienced meteorological officers. In order to make a comparison with reanalysis data, station data have been converted from daily averages to monthly averages, which has the effect of removing high-frequency components from the time series (McGrath and Lynch, 2008).

2.5. Co-location techniques

Spatially, the parameter values in the reanalysis datasets represent the value of that parameter in a grid box centred on the geographic coordinates given in the dataset. This presents a difficulty when comparing the station data with the reanalysis data, because the station data represent a single site within a grid box. Additional complications are introduced by the fact that the grid sizes in ERA-40, ERA-Interim and NNRP-1 are different, the grids in the three reanalysis datasets are not coincident and the location of a station in the grid box can vary from the centre of the grid box to its edge. This is illustrated in Figure 1 where the grids used by ERA-40, ERA-Interim

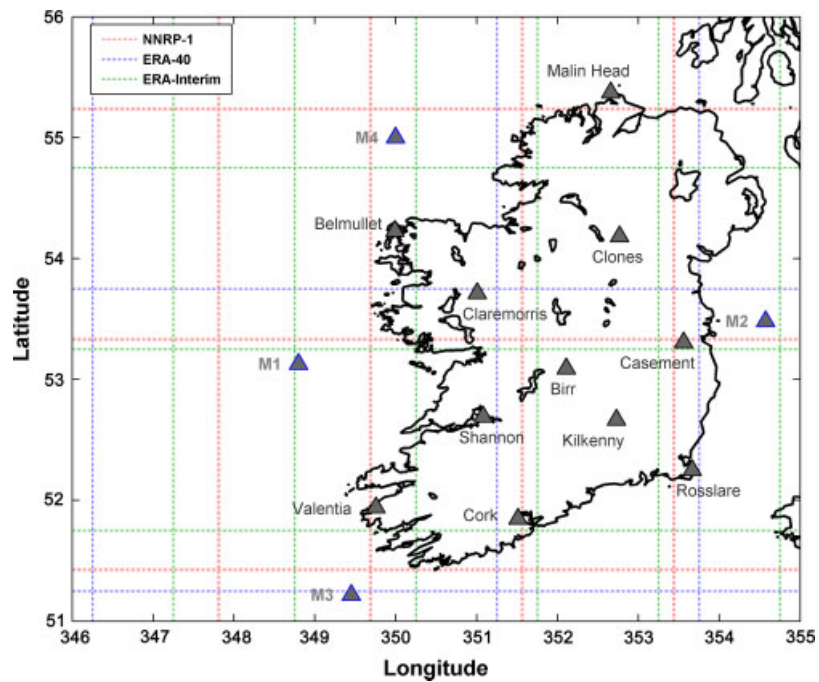


Figure 1. Map of Ireland showing the location of the 11 meteorological stations and the marine buoys M1, M2, M3 and M4. The map is overlaid with grids for NNRP-1 (red – line), ERA-40 (blue – lines) and ERA-Interim (green – line). This figure is available in colour online at wileyonlinelibrary.com/journal/joc

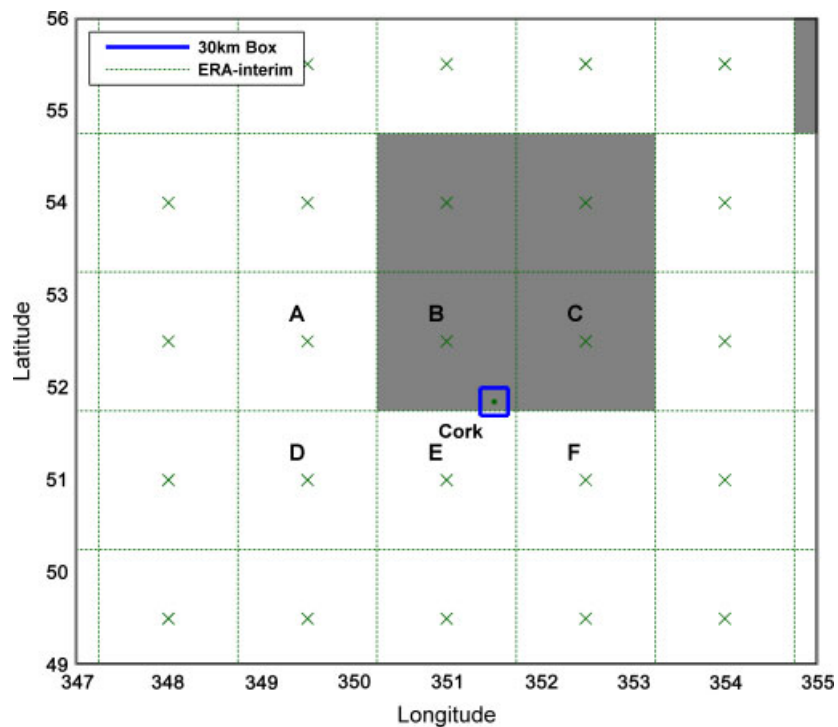


Figure 2. Illustration of the grid boxes used in the assignment of a reanalysis value for comparison with that observed at a given station. The number of grid boxes used depends on the co-location method employed. A grid system is shown with grid boxes labelled A, B, C, D, E and F. The centre of each grid box is marked with a cross, x. A station is represented by a dot and it is surrounded by an imaginary region which intersects the grid boxes B and E. See text for a description of the four co-location methods. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

and NNRP-1 are plotted together with the location of the stations and marine buoys.

Four slightly different approaches were investigated to identify the appropriate reanalysis value for comparison

with the observed values. Figure 2 is used to illustrate the difference in the four approaches. The simplest technique is to compare the value observed at the station with the reanalysis value for the grid box in which the station

lies. In this method, the reanalysis value of grid box B in Figure 2 would be compared with the value observed at the station. This technique will be referred to as method 1.

A second approach – method 2 – was used in which the reanalysis value assigned for comparison with the station value was obtained from the weighted average of the reanalysis values of the four grid boxes whose centres lie closest to the station. The average of the four grid boxes is obtained from the inverse distance weighted average (Stahl *et al.*, 2006). In the case of Figure 2, the weighted average of the grid boxes B, C, E and F was assigned for comparison with the station value.

A third approach, employed by Ruti *et al.* (2008), is to consider a square region of known dimension centred on the station. Ruti *et al.* (2008) chose a 15×15 km region centred on the station and then assigned the average of the reanalysis grid boxes overlapped by the square as the reanalysis value to compare with the station data. In their case, the average was also weighted by the inverse distance method. In the present study, a region of dimension 30×30 km was chosen because it allows a reanalysis value different to those obtained from either of the previous two methods to be assigned to at least two stations for both NNRP-1 (Casement, Cork, Malin Head and Valentia) and ERA-40 (Claremorris and Rosslare). In the case of Figure 2, the reanalysis value assigned for comparison with the observed value was the weighted average of grid boxes B and E. This technique will be referred to as method 3.

Method 4 determines a value for the reanalysis from the land grid points only in the four nearest grid boxes. A grid point is classified as a land point based on the reanalysis land-sea mask. The land points are averaged using the same inverse distance weighting used in methods 2 and 3. For example, the reanalysis value for the station in Figure 2 is determined from a weighted average of grid boxes B and C only, as these are the only 'land' grid boxes of B, C, E and F.

The co-location methods were investigated by examining their effect on the linear correlation coefficient and the root mean squared error (RMSE). Figure 3(a) and (b) shows a plot of the Pearson correlation coefficients and the RMSE at each of the 11 stations for all four methods over the period 1989–2001 for NNRP-1. Figure 3(b)–(e) shows the same plots for the correlation coefficients and the RMSE for ERA-40 and ERA-Interim, respectively. The correlation coefficient of NNRP-1 data at Birr (S2) is identical for all four methods; this is also observed in both ERA-40 (Figure 3(c)) and ERA-Interim (Figure 3(e)). The remaining stations show some variations between the four methods, but in general over the three reanalysis datasets no one method is obviously superior. This is further borne out in the comparison of the mean of the 11 correlation coefficients for each method in NNRP-1 (0.986, 0.989, 0.988 and 0.991 for methods 1, 2, 3 and 4, respectively), ERA-40 (0.991, 0.993, 0.991 and 0.993 for methods 1, 2, 3 and 4, respectively) and ERA-Interim (0.994, 0.996, 0.994 and 0.995

for methods 1, 2, 3 and 4, respectively); in the case of NNRP-1, the mean correlation coefficients differ by less than 0.01 but this difference is not statistically significant (at the 95% confidence level) and there is even less difference in the mean correlation coefficients for ERA-40 and ERA-Interim. Analysis of the four methods using the RMSE also shows that it is difficult to determine which method is obviously superior.

Based on these results, it was decided to proceed with only one of the co-location methods. Method 1 was rejected on the basis that it was the least representative of the four methods because a number of the stations are on the edge of grid boxes. Method 2 was adopted for subsequent work in preference to methods 3 and 4 because it contained the least assumptions of the three methods.

3. Results and discussion

Figure 3(a), (d) and (f) shows that ERA-Interim produces linear correlation coefficients closer to unity than ERA-40 and NNRP-1. This indicates a stronger linear relationship with observations by ERA-Interim than either ERA-40 or NNRP-1. A scatter plot of reanalysis data *versus* observational data enables further investigation of the linear relationship between reanalysis data and observations. A least-squares line fitted to this data could produce lines with slope values equal to, less than or greater than 1.0. Slope values equal to 1.0, in conjunction with an intercept value of 0.0 imply that the reanalysis is in perfect agreement with the observations. If the slope is less than 1.0, then the reanalysis models warmer winters and/or colder summers than the observations. Slope values greater than 1.0 indicate that the reanalysis values are colder in winter and/or warmer in summer than the observations.

Figure 4 is a scatter plot of the ERA-40, ERA-Interim and NNRP-1 datasets *versus* the observed data at Valentia (10.22°W , 51.93°N) for the period 1989–2001. A least squares linear fit to each of the three datasets is also shown. Plots similar to Figure 4 were made for the other 10 synoptic stations shown in Figure 1. The best fit slopes and the standard errors for the 11 stations for ERA-40, ERA-Interim and NNRP-1 are plotted in Figure 5 for the period 1989–2001 along with the slope values for ERA-Interim obtained using co-location method 1. Differences between the slopes for ERA-Interim obtained using co-location methods 1 and 2 are less than 0.05; indicating that the choice of co-location method does not have a significant influence on the results. It is clear that the three reanalysis datasets are in good agreement with observational data; eight of the stations show a slope greater than 0.85 for all three reanalyses.

Figure 5 shows that the slopes of the stations are predominantly below 1.0 with the only exceptions being Malin Head (S8) in the case of ERA-40 and Rosslare (S9) in the case of ERA-Interim as the only exception. This can be understood from the scatter plot in Figure 4 (which is typical of the stations) in which all three reanalysis

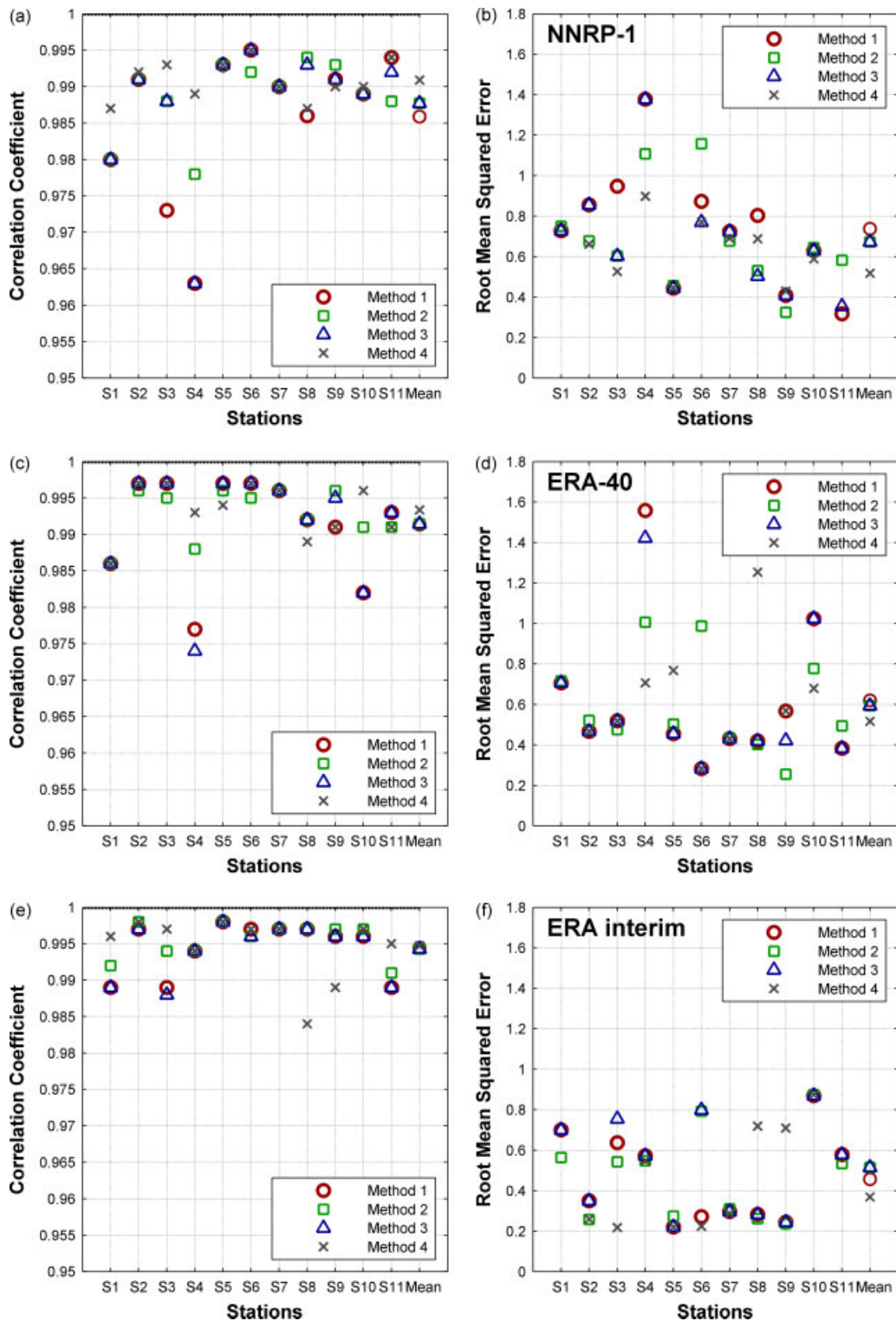


Figure 3. (a) Correlation coefficients for the 11 stations between the observations and NNRP-1 data obtained using the four different co-location methods for the period 1989–2000. (b) Root mean square error for the 11 stations for NNRP-1 data obtained using the four different co-location methods for the period 1989–2000. (c) and (d) same as (a) and (b), respectively, except with ERA-40 data instead of NNRP-1 data. (e) and (f) same as (a) and (b), respectively, except with ERA-Interim data. The 11 stations from S1 to S11 are Belmullet, Birr, Casement, Claremorris, Clones, Cork, Kilkenny, Malin Head, Rosslare, Shannon and Valentia, respectively. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

datasets show warmer values than the observations for winter temperatures. Figure 6(a)–(d) shows plots of the mean monthly temperature over the period 1989–2001 at four representative stations (Birr, Casement, Claremorris

and Valentia) for the ERA-40, ERA-Interim, NNRP-1 and observational data. Each reanalysis dataset accurately captures the mean monthly temperature in the summer months but shows poorer performance in the winter

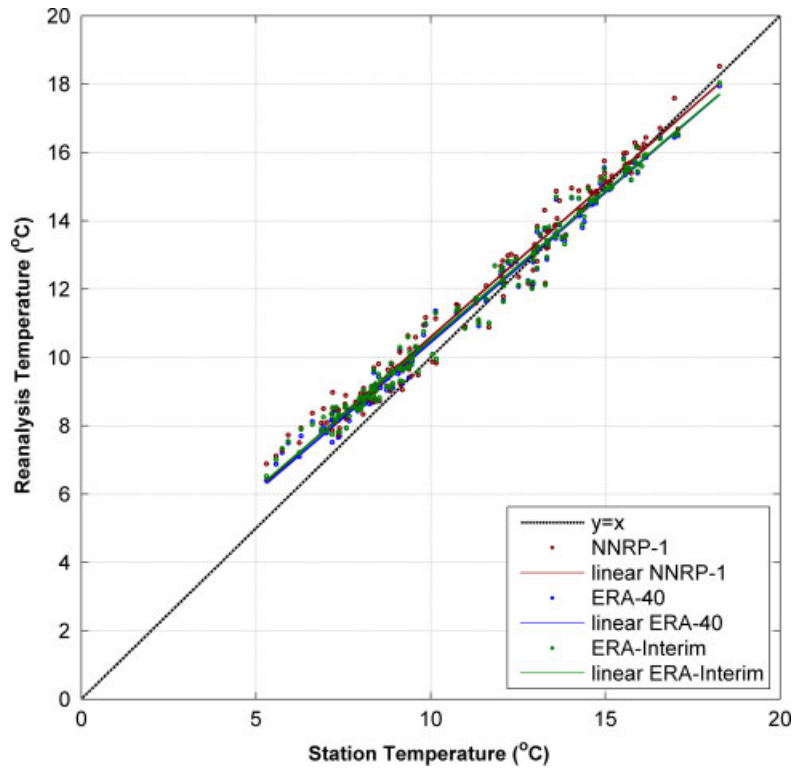


Figure 4. Scatter plot of NNRP-1 (red), ERA-40 (blue) and ERA-Interim (green) reanalysis data assigned using method 2 *versus* observed values for 1989–2001 at Valentia. In each case, the line is the least squares fit of the reanalysis data to the observations. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

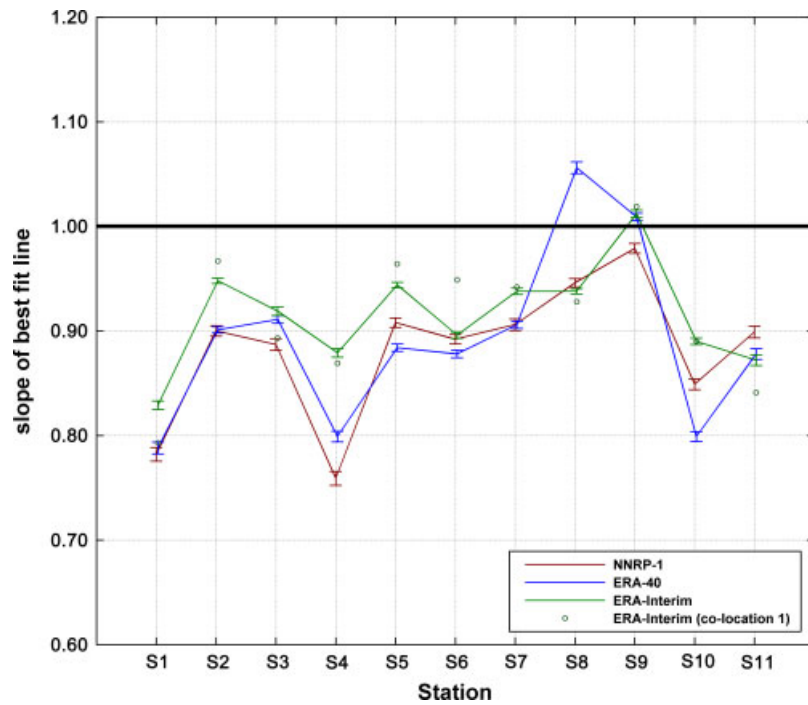


Figure 5. Plot of the slope and standard error of the least squares linear fit of the NNRP-1, ERA-40 and ERA-Interim data to the observed data (determined using co-location method 2 described in the text) for the time period 1989–2000. The same is also shown for ERA-Interim data obtained using co-location method 1 (green open circles). The order of the stations on the x -axis is the same as for Figure 3. In the case of each station, the points are joined by straight lines as an aid to the eye. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

months. This pattern was observed at 10 of the 11 stations. Malin Head, shown in Figure 6(e), deviates from this trend; ERA-40 produces summer values very close to

the observations but models colder winter temperatures which contrasts with ERA-40 values for the other 10 stations. NNRP-1 models colder summer temperatures

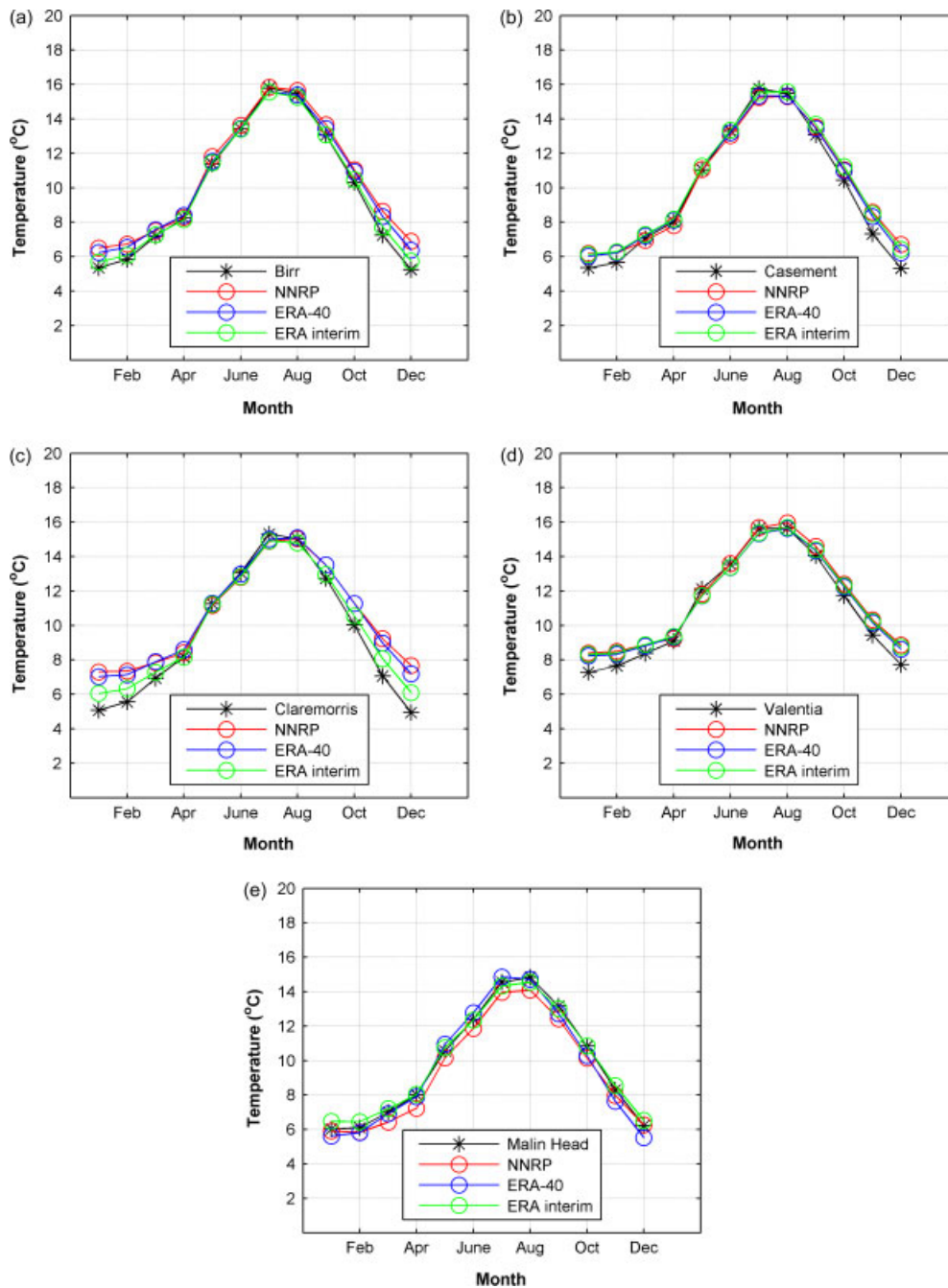


Figure 6. Mean monthly surface air temperatures observed in the period 1989–2001 at Birr, Casement, Claremorris, Valentia and Malin Head. Also plotted in this figure are the corresponding values from NNRP-1 (red), ERA-40 (blue) and ERA-Interim (green) assigned using co-location method 2 described in the text. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

and colder winter temperatures at this station. The temperature bias is greater in summer which causes the best fit slope to be less than unity in agreement with the other 10 stations. ERA-Interim values at Malin Head also produce colder summer temperatures but models warmer winter temperatures producing the typical winter bias present at all stations.

It is difficult to determine from the above results which of NNRP-1, ERA-40 or ERA-Interim is the superior reanalysis product. Prompted by the discrepancies between the temperatures in the reanalyses and observations for the winter months, summary statistics were

examined for both winter (December, January, and February) and summer (June, July and August) seasons for the four stations in Figure 6(a)–(d). These stations were chosen as they are representative of the geographical variation of weather conditions in Ireland; Malin Head was not included in this analysis as it deviates from the typical trend. The location (mean), spread (mean absolute deviation), symmetry (skewness) and kurtosis of the four stations are presented in Figure 7. This figure shows clearly the difference between ERA-40, ERA-Interim and NNRP-1 in their ability to reproduce the observed temperatures. There is little to distinguish between them in

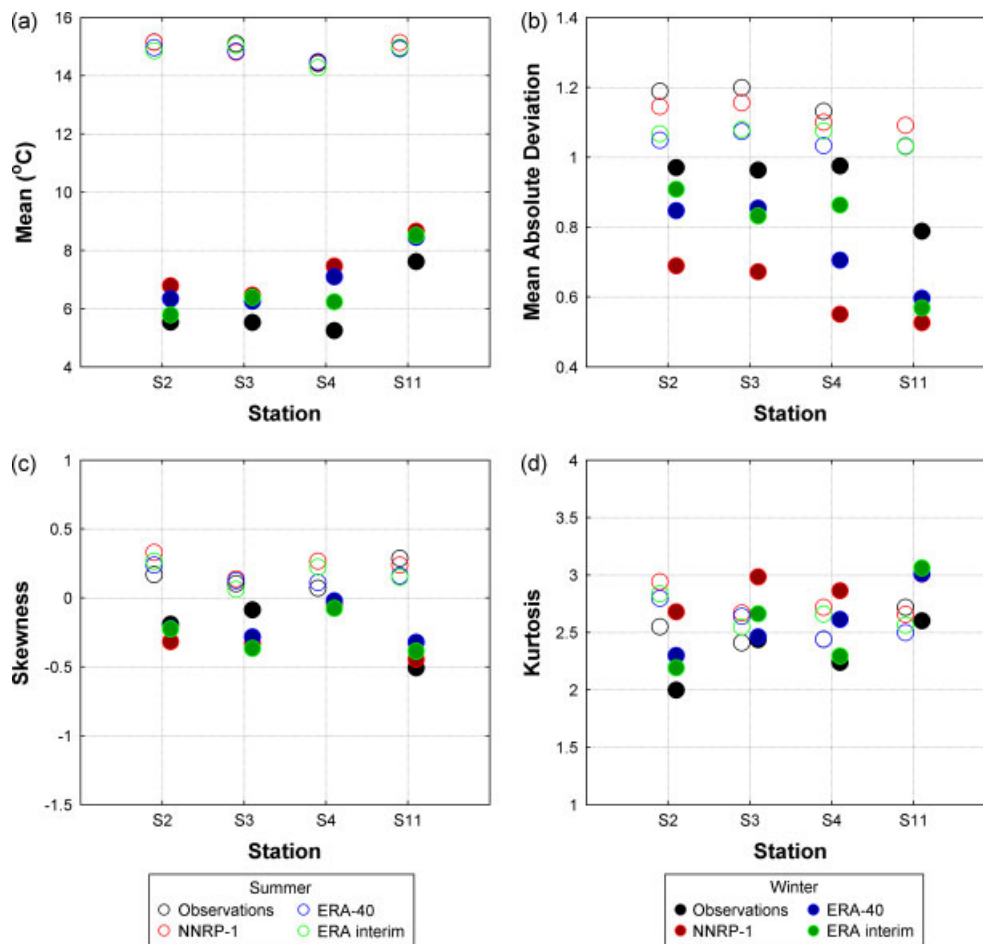


Figure 7. Summary statistics of the observational data, the NNRP-1 data, ERA-40 data and ERA-Interim data for the winter (coloured dots) and summer (open circles) season at four stations (Birr, Casement, Claremorris, Valencia). This figure is available in colour online at wileyonlinelibrary.com/journal/joc

the summer season as shown by the open circles which are very close to the observations. However, in the winter season (filled circles in Figure 7), it is clear that ERA-Interim is consistently closer to the observations than either ERA-40 or NNRP-1. This pattern is reproduced in the skewness and it is shown more clearly in the kurtosis and the mean absolute deviation. Analysis of the Mean absolute error (MAE) shown in Figure 8 also shows that the reanalyses model the summer temperatures better than the winter temperatures and that ERA-Interim models the temperature better than the other two reanalyses.

There are several differences in the assimilation systems of the three reanalyses any one, or any combination, of which could give rise to the above result. One obvious difference is that NNRP-1 data use a horizontal grid of approximately 200 km, whereas ERA-40 and ERA-Interim data are on grids of 250 and 150 km, respectively. It is also noted that the assimilation system of ERA-40 and ERA-Interim uses temperature observations from synoptic weather stations, whereas NNRP-1 does not. Other differences include the method in which each reanalysis produces the temperature at 2 m. NNRP-1 derives the surface air temperature at 2 m from the data

produced by the assimilation system, whereas ERA-40 produces it as part of its assimilation system.

The influence of the grid type and size on the results reported above was examined by transforming the NNRP-1 data and the ERA-Interim data onto the same grid as ERA-40 using three different interpolation methods: linear, cubic and nearest neighbour. Figure 9 shows plots similar to Figure 5 with the slopes of the best fit lines to the NNRP-1 data and ERA-Interim on the larger grid for each of the three interpolation methods examined.

The largest changes occur when the nearest neighbour interpolation method is used. This is not surprising as it is the simplest of the three methods used and it does not make use of all the available information. It is interesting to note that the nearest neighbour method produces the poorest values for ERA-Interim; moving from a smaller grid size to a larger grid size results in a greater loss of data.

Both linear and cubic interpolations use more of the available data than the nearest neighbour method and are a better representation of the ERA-Interim and NNRP-1 data. Figure 9 shows that when these two interpolation methods are used to transform ERA-Interim data and NNRP-1 data onto the larger grid, there is

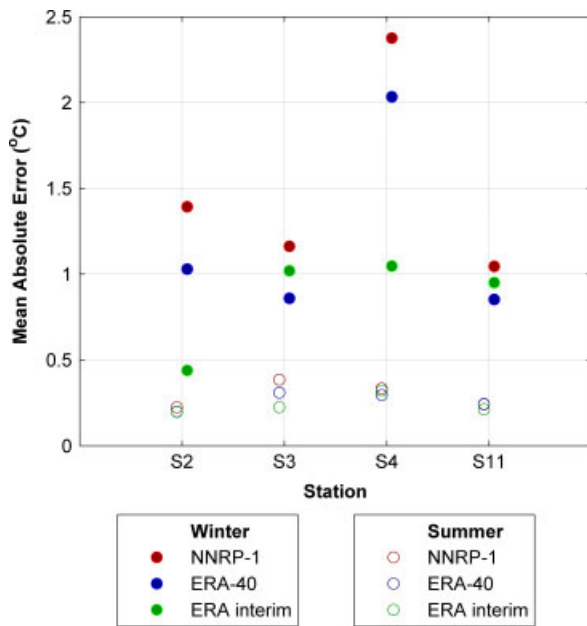


Figure 8. MAE for NNRP-1, ERA-40 and ERA-Interim for the winter (filled circles) and summer (open circles) season at four stations (Birr, Casement, Claremorris, Valentia). This figure is available in colour online at wileyonlinelibrary.com/journal/joc

less variability in the slopes at each station. Despite the improved consistency between the three reanalyses, the transformation of the grids does not significantly affect the overall trend of the slopes (they are consistently less than unity) and ERA-Interim is still marginally closer to unity than the other two reanalyses. Simmons *et al.* (2004) reported a similar finding, i.e. the transformation does not significantly affect the results when they used linear interpolation to transform reanalysis data onto the $5^\circ \times 5^\circ$ resolution grid used by CRUTEMv2 data.

4. Comparison of NNRP-1 and ERA-Interim with marine data

The warmer winter values of the reanalyses at the land stations prompted an investigation of the ability of the reanalyses to reproduce the temperature at marine locations since the models treat marine and land surfaces differently. The time series of hourly temperature observations was obtained for four marine buoys from the Marine Institute who maintains these buoys in cooperation with Met Éireann. The locations of the marine buoys (M1, M2, M3 and M4) are shown on the map in Figure 1.

The observational data from the marine buoys and the four land stations shown in Figure 7 are compared with NNRP-1 and ERA-Interim data; ERA-40 data are not included as its dataset terminates just after the deployment of the marine buoys. The study is restricted to the period 2001–2005 because of the lack of marine data prior to 2001 and the availability of the station dataset only until 2005. It should be noted that the four marine buoys began operation on different dates, e.g. M3 began in mid-2003 and M4 in mid-2004 and as a result, some marine buoys have less data points than others.

Figure 10 shows the slopes of the best fit lines and their errors for the above comparison. It is clear from this figure that the slopes of the marine buoys are consistently greater than unity while the slopes for the land stations are less than unity. The average values for the NNRP-1 and ERA-Interim marine slopes are, respectively, 1.09 and 1.10 compared with the average value of the NNRP-1 and ERA-Interim slopes for the four land stations of 0.87 and 0.89, respectively. As the length of the data series used in this analysis is rather short (5 years), it was important to investigate whether this might be the source of the difference between the marine and land data. Figure 11 is a plot of the slopes and errors of the best fit line to station data at Casement for a moving 5-year window starting in 1979 for NNRP-1 data and 1989 for ERA-Interim. It shows that, apart from increasing the uncertainty, comparing NNRP-1 and ERA-Interim data with observational data over shorter periods does not produce large variations in the slopes (a similar plot is included for ERA-40 for completeness). This indicates that the difference between the results for the marine data and the land data is unlikely due to the shorter data length.

Figure 6 showed that the reanalyses produced more accurate temperature values in summer than winter by examining the mean monthly temperature for the synoptic stations over a 13-year period. Results of a similar analysis for the marine data over the period 2001–2005 are shown in Figure 12. Unlike the land stations, two of the four marine buoys, M1 and M3, show the same degree of warming and cooling in the observations and reanalyses for all seasons and a third marine buoy (M4) also shows this for NNRP-1 and the observations. The high quality of these reanalysis values is also evident in Figure 10, where the two slopes closest to unity are those of the same two marine buoys. In the case of the marine buoy M2, which is located in the Irish Sea, both ERA-Interim and NNRP-1 model slightly warmer summer temperatures than those recorded at the buoy but the trend observed in the land stations of warmer NNRP-1 winters is reversed.

In summary, the investigation of the ability of ERA-Interim and NNRP-1 to reproduce the air temperature at marine locations showed that the slopes of the least-squares line to scatter plots of reanalysis data with observations at the marine buoys are consistently greater than unity whereas the land stations have slopes less than unity. We speculate that the difference in the behaviour of the slopes for marine and land locations may be due to the different treatment of land and sea surfaces (e.g. surface albedo, surface evaporation and specific heat capacity) in the reanalyses or that other atmospheric variables in the reanalysis, such as wind and precipitation, may contribute to the differences observed.

5. Conclusions

The ability of ERA-40, ERA-Interim and NNRP-1 to reproduce observed surface air temperatures at 11 synoptic weather stations in Ireland over the period 1989–2001

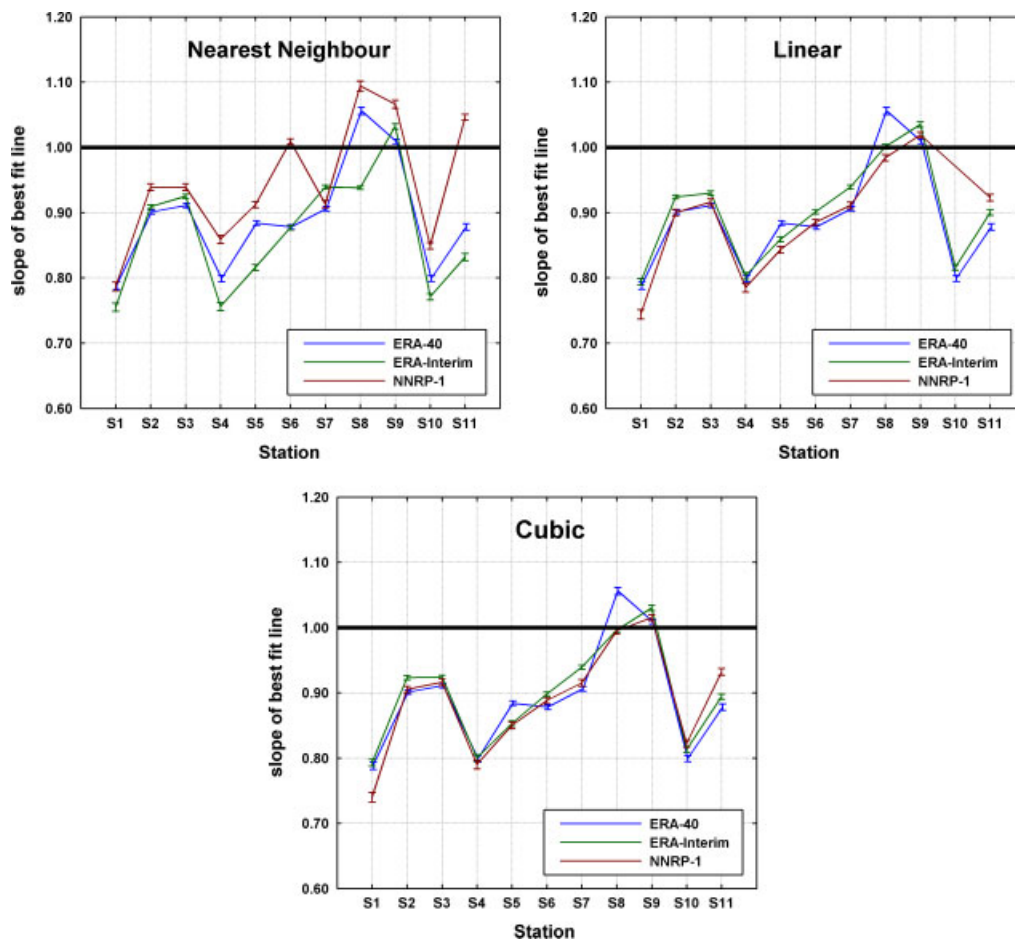


Figure 9. (a) Slopes of least squares line to ERA-40, ERA-Interim and NNRP-1 data. The ERA-Interim and NNRP-1 data have been regridded to the ERA-40 grid using nearest neighbour interpolation. (b) Same as (a) but ERA-Interim and NNRP-1 have been regridded using linear interpolation. (c) Same as (a) but ERA-Interim and NNRP-1 have been regridded using cubic interpolation. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

has been investigated. All three reanalysis datasets show a best fit line to the observational data with slopes greater than 0.85 for 8 of the 11 stations. It was not clear which reanalysis was optimum based on the slope of the best fit line but it was clear that the slopes were generally less than unity. The mean monthly temperatures over the 1989–2001 period showed that all three reanalysis modelled summer temperatures accurately but tended to model significantly warmer winters than were observed. Summary statistics for the winter months showed that while all three reanalyses had difficulty reproducing the winter observations, ERA-Interim was marginally better than ERA-40 and NNRP-1.

A comparison of ERA-Interim and NNRP-1 reanalysis data with marine data for the period 2001–2005 showed that the best fit lines tended to have slopes greater than unity and there was a cold winter bias in the reanalyses at each marine buoy particularly for M2 which also showed a warm summer bias; this contrasts to the situation for the land stations which tended to have slopes less than unity for the same period and a warm winter bias.

Based on the above analysis, it was found that ERA-Interim reproduced the statistics of the observed climate

more favourably than either ERA-40 or NNRP-1. Furthermore, it has been shown that the three reanalyses overestimate the surface temperatures in the winter season indicating a warm winter bias in the reanalyses. Studies using reanalysis data over the Irish terrestrial domain should account for this bias. It appears that the winter bias is not present at the marine buoys and while the limited extent of the data series (2001–2005) makes this a tentative finding, a possible explanation for the result may be found in the difference between the treatment of land and sea surfaces in the reanalysis datasets. This result will be further investigated in future work through the comparison of the reanalyses with satellite data, which will provide greater spatial coverage over land and sea.

Acknowledgements

This study was funded under the Programme for Research in Third Level Institutions (PRTLII) administered by the Irish Higher Education Authority under the National Development Plan and with partial support from the European Regional Development Fund. The authors are very grateful to Met Éireann for

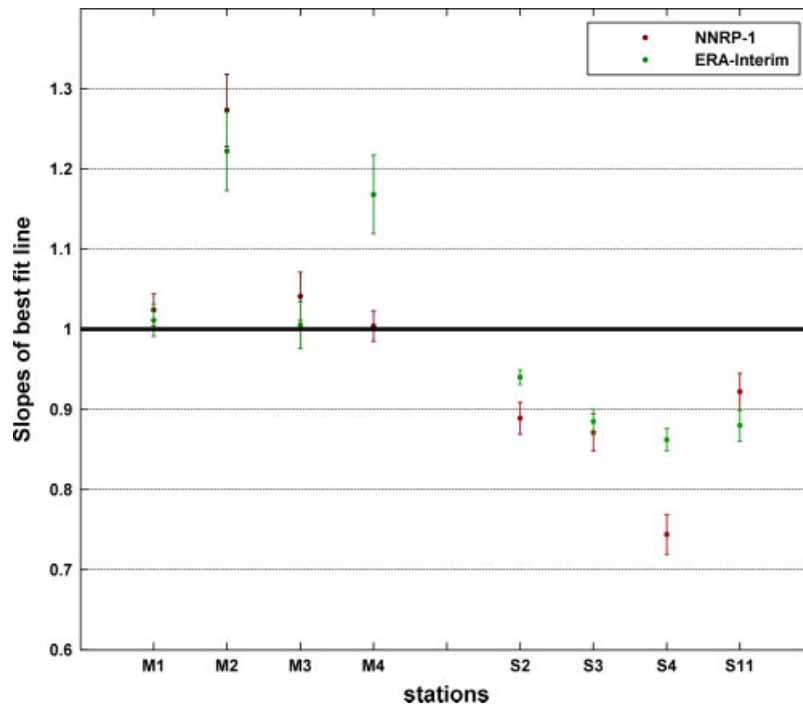


Figure 10. The slopes of the best fit lines for NNRP-1 data and ERA-Interim data to the observational data over the time period 2001–2005. The first four points correspond to the marine buoys M1, M2, M3 and M4 and the remaining four stations S2, S3, S4 and S11 correspond to Birr, Casement, Claremorris and Valentia, respectively. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

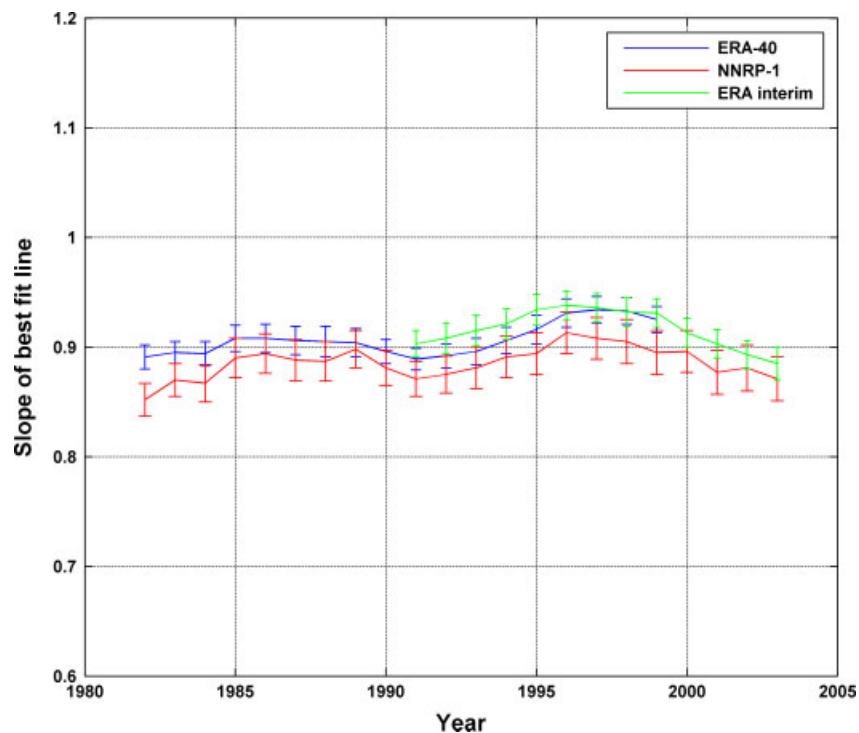


Figure 11. The slopes of the NNRP-1 and ERA-Interim best fit lines to observational data at Casement over a sliding 5-year interval during the period 1979–2005 and 1989–2005, respectively. A plot of the corresponding result for ERA-40 data is included for completeness. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

access to their data from synoptic weather stations and the Marine Institute of Ireland for access to marine buoy datasets. ECMWF ERA-40 and ERA-Interim data used in this study have been obtained from the ECMWF data server: <http://data.ecmwf.int/data>.

NCEP Reanalysis Derived data have been provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at <http://www.cdc.noaa.gov/>. A portion of this study was realised through access to NUI Maynooth's high performance computing resources. The authors are

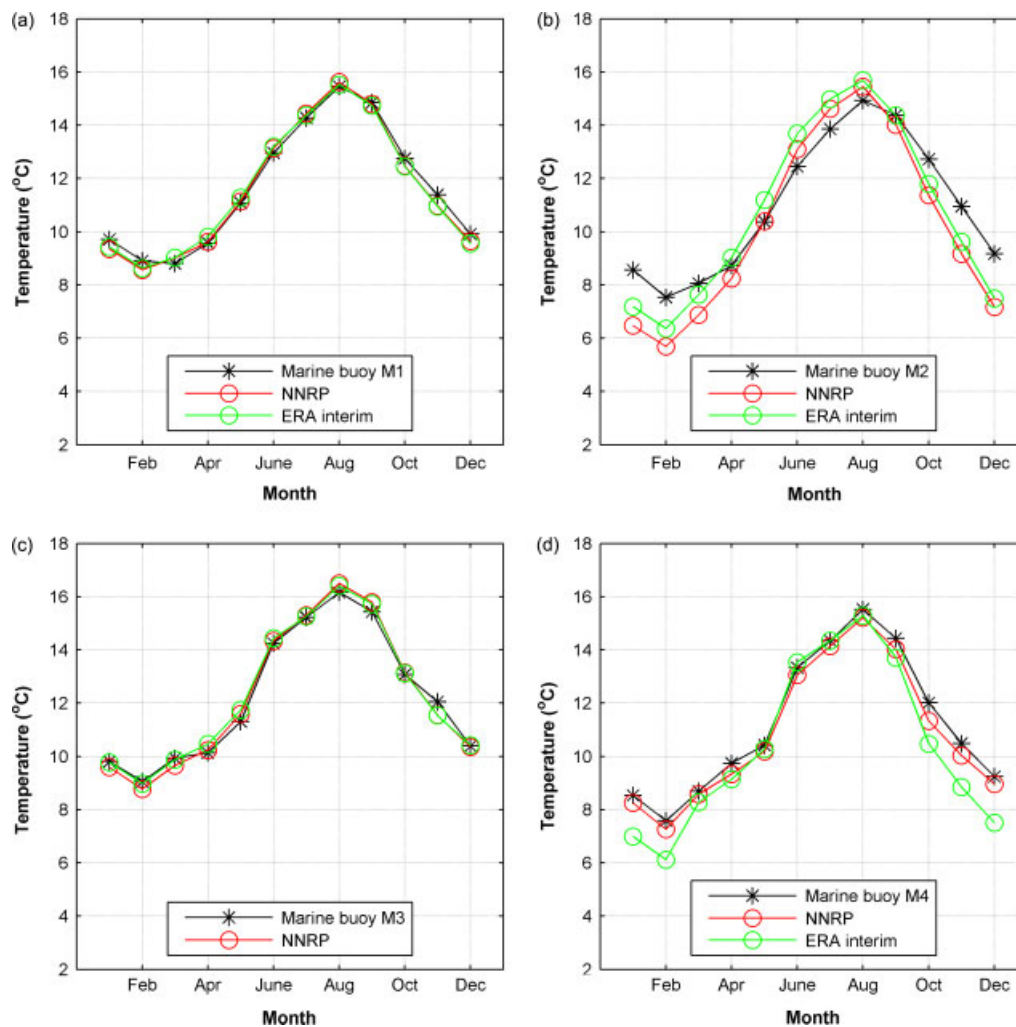


Figure 12. Mean values of the surface air temperature for each month over the period 2001–2005 at the four marine buoys M1, M2, M3 and M4. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

grateful to an anonymous reviewer for many useful comments and suggestions.

References

- Bordi I, Fraedrich K, Petitta M, Sutera A. 2006. Large-scale assessment of drought variability based on NCEP/NCAR and ERA-40 re-analyses. *Water Resources Management* **20**(6): 1573–1650.
- Ciccarelli N, von Hardenberg J, Provenzale A, Ronchi C, Vargiu A, Pelosini R. 2008. Climate variability in north-western Italy during the second half of the 20th century. *Global and Planetary Change* **63**: 185–195.
- Escoffier C, Provost C. 1998. Surface forcing over the South West Atlantic according to NCEP and ECMWF re-analyses over the period 1979–1990. *Physics and Chemistry of the Earth* **23**(5–6): 537–542.
- Fealy R, Sweeney J. 2007. Statistical downscaling of precipitation for a selection of sites in Ireland employing a generalised linear modelling approach. *International Journal of Climatology* **27**: 2083–2094.
- Gleiser H, Thejll P, Stendel M, Kaas E, Machenhauer B. 2005. Solar signals in Tropospheric re-analysis data: comparing NCEP/NCAR and ERA-40. *Journal of Atmospheric and Solar-Terrestrial Physics* **67**: 785–791.
- Henfridsson U, Neimane V, Strand K, Kapper R, Bernhoff H, Danielsson O, Leijon M, Sundberg J, Thorburn K, Ericsson E, Bergman K. 2007. Wave potential in the Baltic Sea and the Danish part of the North Sea, with reflections on the Skagerrak. *Renewable Energy* **32**: 2069–2084.
- Jones PD, Moberg A. 2003. Hemispheric and large-scale surface air temperature variations: An extensive revision and update to 2001. *Journal of Climate* **16**: 206–223.
- Kalnay E. 2007. *Atmospheric Modeling, Data Assimilation and Predictability*, 4th edn. Cambridge University Press: Cambridge.
- Kalnay E, Cai M. 2003. Impact of urbanization and land-use change on climate. *Nature* **423**: 528–531.
- Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woolen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo K, Ropelewski C, Wang J, Leetmaa A, Reynolds R, Jenne R, Joseph D. 1996. The NCEP/NCAR 40-year re-analysis project. *Bulletin of the American Meteorological Society* **77**: 437–471.
- Khan MJ, Iqbal MT. 2004. Wind energy resource map of Newfoundland. *Renewable Energy* **29**: 1211–1221.
- Kistler R, Kalnay E, Collins W, Saha S, White G, Wollen J, Chelliah M, Ebisuzaki W, Kanamitsu M, Kousky V, van den Dool H, Jenne R, Fiorino M. 2001. The NCEP-NCAR 50 year reanalysis: Monthly means CD-ROM and documentation. *Bulletin of the American Meteorological Society* **82**: 247–267.
- Li H, Robock A, Liu S, Mo X, Viterbo P. 2004. *Evaluation of Reanalysis Soil Moisture Simulations using updated Chinese Soil Moisture Observations*. ERA-40 Project Report Series 20.
- McGrath R, Lynch P. 2008. *Ireland in a Warmer World: Scientific Predictions of the Irish Climate in the Twenty-first Century*. Community Climate Change Consortium for Ireland.
- Poveda G, Waylen PR, Pulwarty RS. 2006. Annual and inter-annual variability of the present climate in northern South America and southern Mesoamerica. *Paleogeography, Palaeoclimatology, Palaeoecology* **234**: 3–27.

- Ruti PM, Marullo S, D'Ortenzio F, Tremont M. 2008. Comparison of analyzed and measured wind speeds in the perspective of oceanic simulations over the Mediterranean basin: analyses, quickscat and buoy data. *Journal Marine Systems* **70**: 33–48.
- Simmons AJ, Jones PD, da Costa Bechtold V, Beljaars ACM, Kollberg PW, Saarinen S, Uppala SM, Viterbo P, Weldi N. 2004. *Comparison of Trends and Variability in CRU, ERA-40 and NCEP/NCAR Analyses of Monthly-mean Surface Air Temperature*. ERA-40 project report series 18.
- Simmons A, Uppala S, Dee D, Kobayashi S. 2006. ERA-Interim: New ECMWF reanalysis products from 1989 onwards. *ECMWF Newsletter* **110**: 26–35.
- Stahl K, Moore RD, Foyer JA, Asplin MG, McKendry IG. 2006. Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density. *Agricultural and Forest Meteorology* **139**: 224–236.
- Stammerjohn SE, Martinson DG, Smith RC, Iannuzzi RA. 2008. Sea ice in the western Antarctic Peninsula region: Spatio-temporal variability from ecological and climate change perspectives. *Deep Sea Research II* **55**: 2041–2058.
- Uppala SM, Källberg PW, Simmons AJ, Andrea U, da Costa Bechtold V, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, Van de Berg L, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hólm E, Hoskin BJ, Isaken L, Janssen PAEM, Jenne R, McNally AP, Mahfouf J-F, Morcrette J-J, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J. 2005. The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society* **131**: 2961–3012.
- Wang S, McGrath R, Semmlar T, Sweeney C. 2006. Validation of simulated precipitation patterns over Ireland for the period 1961–2000. *International Journal of Climatology* **26**: 251–266.