Creating localised near future weather data for predicting the performance of buildings in the UK

Hu Du[#1], Phil Jones[#2], Simon Lannon[#3]

[#]Welsh School of Architecture, Cardiff University
Bute Building, King Edward VII Avenue, Cardiff, CF10 3NB, United Kingdom

[1]DuH4@cardiff.ac.uk

[2]jonesp@cardiff.ac.uk

[3]lannon@cardiff.ac.uk

Abstract
Predicting near future building performance is an important aspect of demand response control strategies, which was highlighted in the Horizon 2020 societal challenges work programme by the European Commission. Past research shows that the optimisation of energy management with weather forecasting can generate 15-30% savings in most cases. Therefore, it is crucial to develop a method of gathering reliable weather forecast data and applying the forecast data into building performance simulation or building energy management system. Since 2011, the Meteorological Office (Met Office) in the United Kingdom released 3-hourly site-specific forecast data feeds for nearly 6,000 locations in the UK through the Met Office DataPoint in a format that is suitable for web application developers. This provids a great opportunity for building performance simulation professionals to re-use Met Office data for predicting near future building performance. With the freely available high frequency weather forecast data, the aim of this paper is to create localised near future weather data for predicting the performance of buildings in the UK. The project is built on authors' previous research experience in future weather data and building performance modelling. The authors developed a method of automatically gathering forecast weather data for nearly 6,000 locations in the UK. Through the detailed comparison between forecast and observation, authors are confident that the 24-hour forecasts are very close to observations. Therefore, the high resolution forecasts for the significant large number of locations can be used to create 'real' weather data for locations that do not have weather stations.

Keywords - near future weather data; building simulation; performance predication; epw; Met Office

1. Introduction

The United Kingdom (UK) Government has set a challenging target[1] for reducing greenhouse gas emissions 80% by 2050 (from the 1990 baseline). Given that around 45% of emissions come from buildings, there is no other way to meet the target without addressing the performance of buildings. Buildings often do not perform as predicted in term of energy consumption. The most recent study [2] conducted by the Innovate UK's Building Performance Evaluation programme shows that the measured energy ratings (DEC) are often triple the modelled predicted energy ratings (EPC). Minimising the gap between predicted and measured performance is crucial in reducing carbon emission of buildings [3]. The performance gap is not only decided by physical characteristics (fabric and system) of a building, but also influenced by users (occupants and their behaviours), more importantly by external conditions (weather). Therefore reliable and high quality weather data is essential for all building performance analysis including both modelling (forecasting) and monitoring (observation).

Predicting near future building performance is also an important aspect of demand response control strategies, which was highlighted in the Horizon 2020 societal challenges work programme by the European Commission [4]. The past research [5] shows that the optimisation of energy management with weather forecasting can generate 15-30% savings in most cases. Therefore, it is crucial to develop a method of gathering reliable weather forecast data and applying the forecast data into building performance simulation or building energy management system. It is not only to the interests of building performance researchers, but also to smart home industry. More and more real-time building performance monitoring and control products are available on market in recent years, such as Google Nest Thermostat, British Gas home monitor, BRE home monitoring system and Danfoss heating controller. These systems could easily be adapted to optimise building energy consumption with near future weather forecasting. However, there is missing link between building simulation/energy management and weather forecast.

Since 2011, the Meteorological Office (Met Office) in the UK released 3-hourly site-specific forecast data feeds for nearly 6,000 locations in the UK through the Met Office DataPoint service [6] in a format that is suitable for web application developers - Extensible Markup Language (xml). This provided a great opportunity for building performance simulation professionals to re-use the Met Office data for predicting near future building performance. The Met Office DataPoint service provides last 24-hour land and marine observations, lightning strikes frequency map, monthly regional climate anomalies, historic regional climate data and more importantly 3-hourly weather forecast for next 5 days. The last 24-hour observations are based on 140 weather stations located cross the UK (as shown in Fig. 1 left); the forecasts were made for approximately 6,000 sites around the UK (512 in Wales, green dots in Fig. 1 right), which covers population centres, sporting venues and tourist attractions that are interested to people. Forecasts are regularly updated and make use of a range of weather models run on the Met Office supercomputer. Comparing with the traditional 14 locations of CIBES weather data, the 6,000 sites forecasts provide fine resolution in term of geographical distribution of weather data sites. This could reduce the distance between case study building and weather data site from few hundreds kilometres to few kilometres or even less.

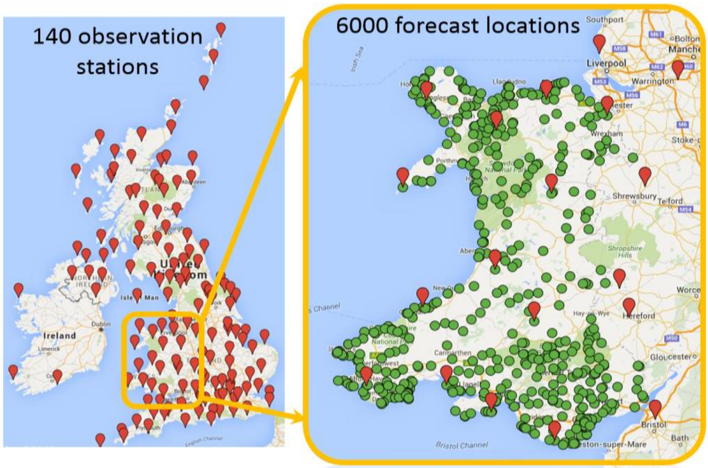
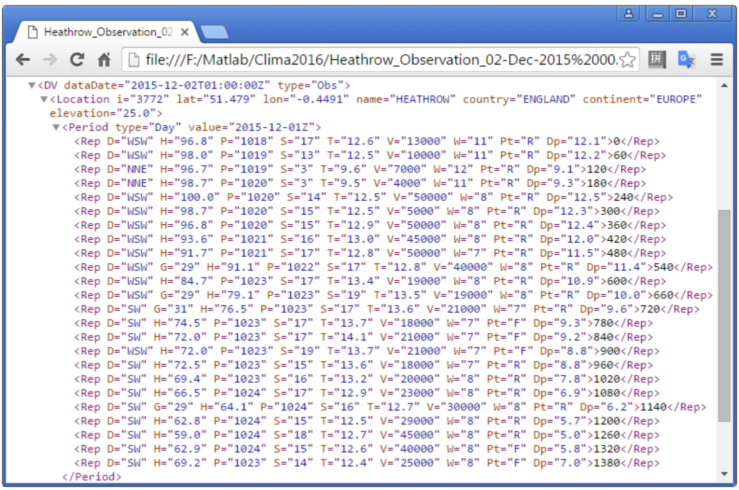


Fig. 1 140 weather stations around the UK for observation (left) and forecasts for approximately 6,000 sites in the UK (512 in Wales, green dots in the right figure)




```xml
<DV dataDate="2015-12-02T01:00:00Z" type="Obs">
 <Location i="3772" lat="51.479" lon="-0.4491" name="HEATHROW" country="ENGLAND" continent="EUROPE"
 elevation="25.0">
  <Period type="Day" value="2015-12-01Z">
    <Rep D="WSW" H="96.8" P="1018" S="17" T="12.6" V="13000" W="11" Pt="R" Dp="12.1">0</Rep>
    <Rep D="WSW" H="98.0" P="1019" S="13" T="12.5" V="10000" W="11" Pt="R" Dp="12.2">60</Rep>
    <Rep D="NNE" H="96.7" P="1019" S="3" T="9.6" V="7000" W="12" Pt="R" Dp="9.1">120</Rep>
    <Rep D="NNE" H="98.7" P="1020" S="3" T="9.5" V="4000" W="11" Pt="R" Dp="9.3">180</Rep>
    <Rep D="WSW" H="100.0" P="1020" S="14" T="12.5" V="50000" W="8" Pt="R" Dp="12.5">240</Rep>
    <Rep D="WSW" H="98.7" P="1020" S="15" T="12.5" V="5000" W="8" Pt="R" Dp="12.3">300</Rep>
    <Rep D="WSW" H="96.8" P="1020" S="15" T="12.9" V="50000" W="8" Pt="R" Dp="12.4">360</Rep>
    <Rep D="WSW" H="93.6" P="1021" S="16" T="13.0" V="45000" W="8" Pt="R" Dp="12.0">420</Rep>
    <Rep D="WSW" H="91.7" P="1021" S="17" T="12.8" V="50000" W="7" Pt="R" Dp="11.5">480</Rep>
    <Rep D="WSW" G="29" H="91.1" P="1022" S="17" T="12.8" V="40000" W="8" Pt="R" Dp="11.4">540</Rep>
    <Rep D="WSW" H="84.7" P="1023" S="17" T="13.4" V="19000" W="8" Pt="R" Dp="10.9">600</Rep>
    <Rep D="WSW" G="29" H="79.1" P="1023" S="19" T="13.5" V="19000" W="8" Pt="R" Dp="10.0">660</Rep>
    <Rep D="SW" G="31" H="76.5" P="1023" S="17" T="13.6" V="21000" W="7" Pt="R" Dp="9.6">720</Rep>
    <Rep D="SW" H="74.5" P="1023" S="17" T="13.7" V="18000" W="7" Pt="F" Dp="9.3">780</Rep>
    <Rep D="SW" H="72.0" P="1023" S="17" T="14.1" V="21000" W="7" Pt="F" Dp="9.2">840</Rep>
    <Rep D="WSW" H="72.0" P="1023" S="19" T="13.7" V="21000" W="7" Pt="F" Dp="8.8">900</Rep>
    <Rep D="SW" H="72.5" P="1023" S="15" T="13.6" V="18000" W="7" Pt="R" Dp="8.8">960</Rep>
    <Rep D="SW" H="69.4" P="1023" S="16" T="13.2" V="20000" W="8" Pt="R" Dp="7.8">1020</Rep>
    <Rep D="SW" H="66.5" P="1024" S="17" T="12.9" V="23000" W="8" Pt="R" Dp="6.9">1080</Rep>
    <Rep D="SW" G="29" H="64.1" P="1024" S="16" T="12.7" V="30000" W="8" Pt="R" Dp="6.2">1140</Rep>
    <Rep D="SW" H="62.8" P="1024" S="15" T="12.5" V="29000" W="8" Pt="R" Dp="5.7">1200</Rep>
    <Rep D="SW" H="59.0" P="1024" S="18" T="12.7" V="45000" W="8" Pt="R" Dp="5.0">1260</Rep>
    <Rep D="SW" H="62.9" P="1024" S="15" T="12.6" V="40000" W="8" Pt="F" Dp="5.8">1320</Rep>
    <Rep D="SW" H="69.2" P="1023" S="14" T="12.4" V="25000" W="8" Pt="F" Dp="7.0">1380</Rep>
  </Period>
```


Fig. 2 A sample XML file showing observed hourly weather data of Heathrow Airport on 1st December 2015

The Extensible Markup Language (xml) file generated by the Met Office DataPoint service is following the defined rules for encoding documents in a format which is both human-readable and machine-readable.

It is defined by the World Wide Web Consortium's XML 1.0 Specification. As an example, the xml file (Fig. 2) shows the last 24-hour observed weather data of Heathrow Airport on 1st December 2015. It includes name, latitude (lat), longitude (lon), elevation, wind direction (D), relative humidity (H), wind gust (G), sea level pressure (P), wind speed (S), temperature (T), visibility (V), weather type (W), pressure tendency (Pt) and dew point temperature (Dp). The symbols (in the brackets above) are shown in orange letters in Fig. 2. The value of each parameters are shown in blue and time series are shown in black.

2. Aims and methodology

With the high frequency weather forecast data freely available for approximately 6,000 sites in the UK from the Met Office DataPoint service, the overarching research aim of this study is to create localised near future weather data for predicting the performance of buildings in the UK. The project is built on authors' previous research experience in future weather data [7] [8] and building performance modelling [9]. To achieve this aim, the research approach is:

(i) Developing a method of collecting near future weather forecast for nearly 6000 locations in the UK. Authors wrote a Matlab script to automatically acquire 3-hourly weather forecast data through the Met Office DataPoint service five times a day at midnight, 2am, 4am, 6am and 12pm starting from 1st December 2015 until the time while the paper was prepared (16th January 2016). Due to the Met Office's fair use policy (no more than 5,000 data requests per day), authors only recorded the forecasts for 140 locations that have observation data. This is for the comparison study presented in later section. The forecast is updating every hour, therefore it is important to log the data at a fixed cycle to provide convenience for data capturing in later stage. Every day 700 xml files (140*5) containing 3-hourly forecast for next 5 days were created and stored on a local computer. It takes 2-3 minutes to run the 140 requests at the defined time. The file size of each xml file is ranging from 1KB to 5KB depending on the amount of data contained.

(ii) Applying the similar method to collect hourly weather observation data from 140 weather stations in the UK. The similar Matlab script was created to request observed data for the 140 locations at midnight, 2am, 4am, 6am and 12pm every day. This results another 700 xml files stored on the local computer. The file size of observation data is ranging from 1KB to 3KB depending on the number of missing data or breakdown of weather stations. The xml files conation the last 24-hour hourly weather data immediately before the requests were made. Therefore it is necessary to acquire more than once a day to avoid any missing data.

(iii) Capturing forecast and observation data from individual xml files and comparing the forecasts with observations to understand the confidence of near future weather forecast in the UK. For each location, the authors wrote a Matlab script to capture a particular parameter (such as temperature presented in this paper) from xml files (generated in (i) and (ii) steps over the logging period. A single csv file was generated to contain time series observation data over the whole logging period (from 1st December 2015 – 16th January 2016). Five csv files containing forecast data were also generated by the Matlab script. The deference of the five files is explained in the later section. Because of the breakdowns of weather stations over short period, interpolations were conducted to fill the missing data. For the four locations and logging period studied in this paper, the maximum duration of missing data is less than 6 hours. This should not have significant impact on further analysis. The comparison between observation and forecast was initially conducted for the a particular site (Cardiff) over 5 days, then it was extended to a longer term (47 days) and to four sites cross the UK (London, Cardiff, Edinburgh and Belfast). The results of comparison were presented in the following section. Please note that the forecast data were at 3-hour interval and observation data were at 1-hour interval, therefore the comparison were made at 3-hour interval where there is forecast data.

(iv) Converting the Extensible Markup Language (.xml) weather forecast files into building simulation software readable weather files (such as .epw format). This step involves a detailed comparison between parameters provided by the Met Office DataPoint service and the parameters needed for building performance simulation tools, such as weather variables in epw weather files. The finding and limitations are presented in the later section of this paper. This paper is focused on the analysis of temperature only due to the page limit, the full analysis of other parameters and simulation study will be reported in a journal paper.

3. Results

Comparison of forecast and observation, 5-day data of Cardiff

Because of the rolling update feature of forecast data, 5-day forecasts could be generated each hour based on the latest information of weather model and observation data. Therefore 120 (24*5) sets of forecasts could potentially be requested over 5-day period from the Met Office DataPoint service. To simplify the comparison and reporting, Fig. 3 illustrated the 5-day observation (black line) for Cardiff (St Athan Airport, see details of weather station in Fig. 6) and five forecasts made at midnight of each day. For example, the green line represents the 5-day forecast made at the end of day 0, orange line represents the forecast made at the end of day 1 and red line represents the forecast made at end of day 4. As shown in the figure, the first 1-2 day forecast is reasonable close to observation (e.g. green line follows black line in the first two days), and the forecasts for day 4 and 5 are improving when the forecasts were made close these days, for example, blue, purple and red lines are close to observation in day 5, comparing with green and orange lines which are the forecasts made few days ago.

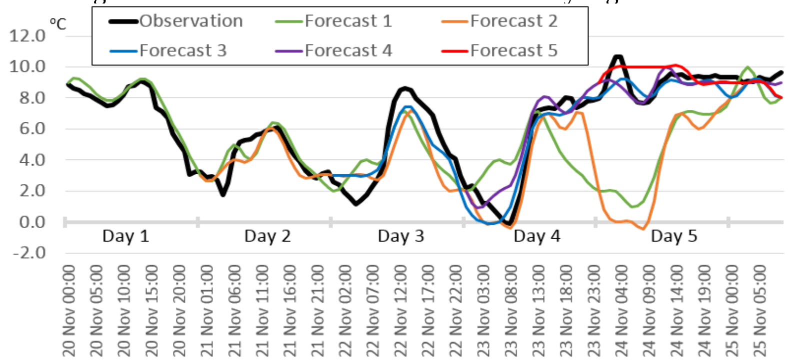


Fig. 3 Temperature forecast and observation, 5-day data of Cardiff (20th -25th Nov 2015)

Comparison of forecast and observation, 47-day data of Cardiff

Following the analysis above, a continue 47-day weather file were constructed using Matlab. It always took the first day of 47 forecasts made at midnight of each day over the logging period (as illustrated in green block in Fig. 4, authors named it as '1 day forecast' weather file). Similarly, 2 day forecast (yellow blocks), 3 day forecast (blue), 4 day forecast (grey) and 5 day forecast (purple) were constructed.

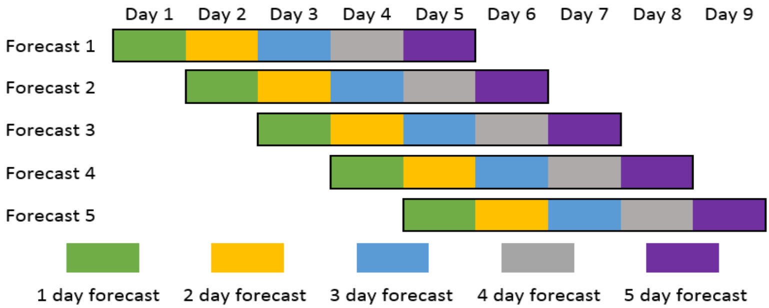


Fig. 4 Construction of forecast weather files

The comparison (Fig. 5) was made between observation (black line in Fig. 5) and 1 day forecast (green line) over the whole logging period from 1st Dec 2015- 16th Jan 2016. It shows that 1 day forecasts are fairly close to observations apart from missing few peaks. The missing data in the forecast files or the 3-hour interval of forecast, which is longer than 1-hour interval observation, could potentially cause this. Similar studies were conducted with other four forecast files as well. Due to the page limits, the only distributions of the differences between 5 forecasts and observation over 47-

days are shown in Fig. 7 (in next page due to large size of the figure). The distributions show that 1 day forecasts are generally closer to observations comparing with others. The mean value (0.18) of the differences for 1 day forecasts is the smallest and same for the standard deviation (0.772). This indicates that weather forecast is more accurate when time is moving forward. For example, the forecast made today is more accurate than the forecast made yesterday. In average, 1 day forecasts are only 0.18 °C away from the observed temperatures, and for the most of time (68% out of 376 forecasts made) the differences between forecasts and real temperatures are less than 0.772 °C. The 5 day forecasts are 0.68 °C below the real temperatures in average, and the standard deviation increases to 1.514 °C.

Fig. 5 Temperature forecast and observation, 47-day data of Cardiff (starting 1st Dec 2015)

Comparison of forecast and observation, 47-day data cross 4 sites

Further analysis were conducted for another three sites cross the UK. Including the previous analysis for Cardiff, the study covers capitals of England, Wales, Scotland and Northern Ireland. The location and details of weather stations are listed in Fig. 6. Same period (1st Dec 2015- 16th Jan 2016) were used for the analysis. The results in Fig. 8 shows that 1 day forecasts provide fairly accurate estimation of temperature with average difference less than 0.2 °C cross all sites. In most of situations, the differences are less than 1 °C.

City	Weather station	Latitude	Longitude	Country
London	Heathrow	51.479	-0.449	England
Cardiff	St Athan Airport	51.405	-3.440	Wales
Edinburgh	Gogarbank	55.928	-3.343	Scotland
Belfast	Belfast Airport	54.664	-6.224	Northern Ireland

Fig. 6 Locations of case study sites

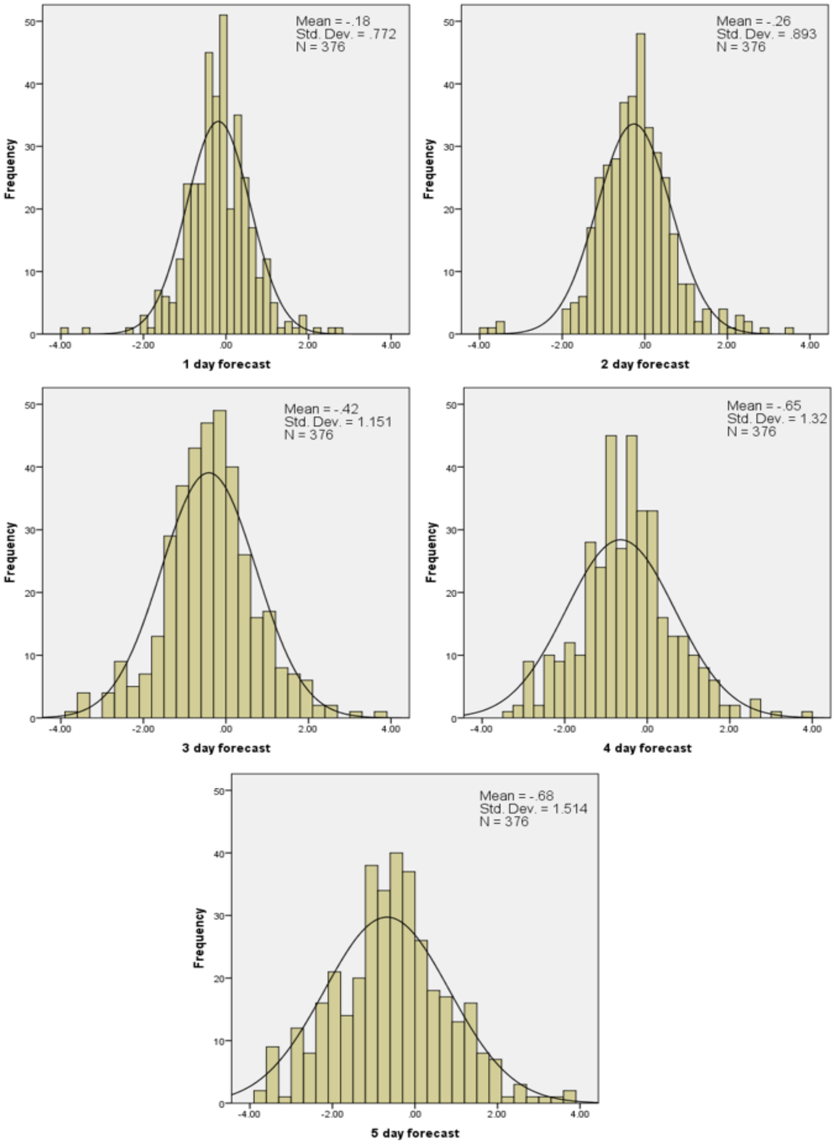


Fig. 7 Distributions of the differences between 5 forecasts and observation over 47-day for Cardiff

Through the detailed comparison between forecast and observation, authors are confident that the 24-hour forecasts are very close to observations. Therefore, the high resolution forecasts for the significant large number of locations can be used to create 'real' weather data for locations that do not have weather stations.

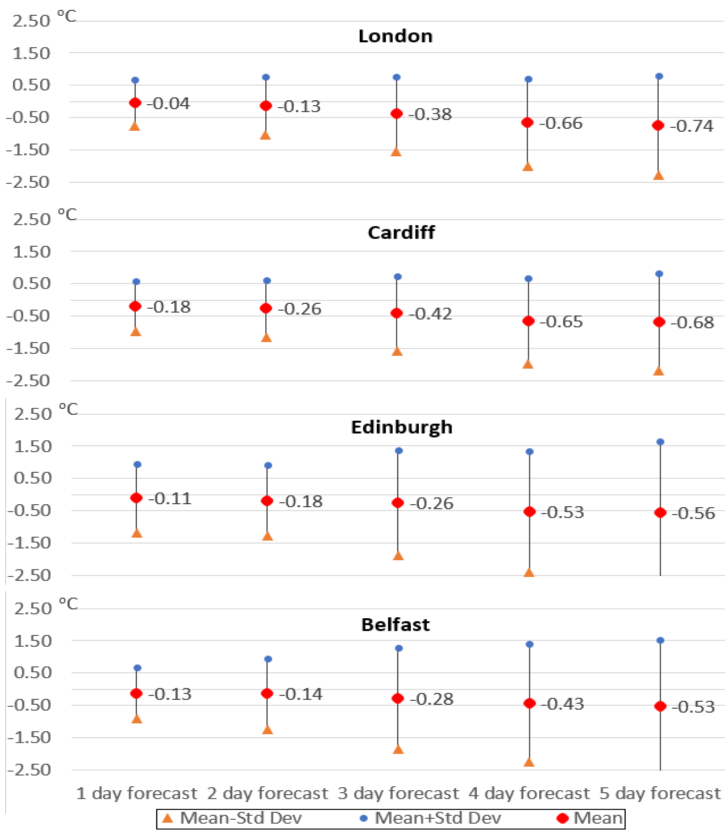


Fig. 8 Means and standard deviations of the differences between forecasts and observations for the four sites

Other parameters needed for building simulation

Further investigation was conducted to compare all parameters provided by the DataPoint service with parameters needed for building simulation. Authors found that the 16 point compass wind direction data need to be converted to degree format for simulation programme. Wet bulb temperature does not appear in the Met Office's data, however it can be calculated from dry bulb temperature, relative humidity and air pressure. The key challenge of creating weather file for building simulation is the absence of solar radiations (both diffuse and global) in Met Office's forecast data. However, with the availability of location, visibility, UV index and other descriptive data, such as weather code, solar radiations could potentially be calculated using solar geometry equations.

4. Main conclusions

The authors developed a solution to re-use the Met Office weather forecast data in building simulation field and energy management system. With the large data sets available for nearly 6000 locations in the UK, it provided a method to create 'real' weather data for locations that do not have weather observation station. The near future 'real' weather data not only can reduce the error in calculation of energy demand, but also help optimise energy management.

Through the comparison of observed temperatures and forecasts, the authors understood the accuracy of near future temperature forecast. The comparison between parameters provided by the DataPoint service and parameters needed for building simulation has also been made to identify the difficulty in implementing energy optimisation with near future weather forecast. With the confidence in temperature forecasting, authors will explore the accuracy of other key parameters for building performance and their impacts on thermal comfort and energy optimisation options.

Acknowledgment

This study is part of the Sêr Cymru ('Stars Wales' in English) programme funded by the Welsh government. The authors thankfully acknowledge the reliable DataPoint service provided by the UK Met Office. The authors also would like to thank W. Falkena at Delft University of Technology and others for sharing the 'xml2struct.m' Matlab function.

References

[1] Great Britain, The Climate Change Act 2008 Chapter 27. 2008, The Stationery Office: London. [2] Palmer, J. and P. Armitage, Early findings from non-domestic projects, Innovate UK, Editor. 2014. [3] de Wilde, P., The gap between predicted and measured energy performance of buildings: A framework for investigation. Automation in Construction, 2014. 41: p. 40-49. [4] European Commission, Horizon 2020: EE-06-2015 Demand response in blocks of buildings. 2015. [5] Lazos, D., A.B. Sproul, and M. Kay, Optimisation of energy management in commercial buildings with weather forecasting inputs: A review. Renewable and Sustainable Energy Reviews, 2014. 39: p. 587-603. [6] The Met Office. The Met Office DataPoint. 2016; Available from: http://www.Met Office.gov.uk/datapoint. [7] Du, H., C.P. Underwood, and J.S. Edge, Generating design reference years from the UKCP09 projections and their application to future air-conditioning loads. Building Services Engineering Research and Technology, 2012. 33(1): p. 63-79. [8] Du, H., C.P. Underwood, and J.S. Edge, Generating test reference years from the UKCP09 projections and their application in building energy simulations. Building Services Engineering Research and Technology, 2012. 33(4): p. 387-406. [9] Du, H., J. Edge, and C. Underwood. Modelling the impacts of new UK future weather data on a school building. in Building Simulation 2011. 2011. Sydney, Australia: International Building Performance Simulation Association