Atmospheric Research 164–165 (2015) 27–36

Contents lists available at ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmos

Bayesian Inference aided analog downscaling for near-surface winds in complex terrain

Alon Manor [1], Sigalit Berkovic

Israel Institute for Biological Research, P.O.B. 19, Ness-Ziona, Israel

ARTICLE INFO

Article history:
Received 11 January 2015
Received in revised form 13 April 2015
Accepted 20 April 2015
Available online 28 April 2015

Keywords:
Statistical downscaling
Numerical weather prediction
Complex terrain

ABSTRACT

Assessing atmospheric boundary layer flows in complex terrain for short-range real-time applications demands fast and reliable downscaling from coarser-resolution meteorological data to the relevant scale. An ideal statistical downscaling numerical experiment was performed for surface winds above complex terrain in Israel's northern Negev desert region. Dynamical downscaling have been performed by the WRF model to create a historical database by the following two sets. The first set used 5 nested domains from 40.5 km to 0.5 km. The second set used 3 nested domains ranging from 40.5 km to 4.5 km. The 4.5 km data (stage 2) was defined as predictors while data on 0.5 km (stage 1) served as predictands for statistical downscaling. Two statistical downscaling algorithms: minimal distance analog and a Bayesian inference aided analog (hereafter Bayesian algorithm) were tested by the above data. Unlike most analog algorithms, the Bayesian algorithm refers to the probability to get the best analog instead of the minimal differences between predictands. The comparison of the two algorithms shows that the Bayesian approach yields improved results. The Bayesian algorithm reproduces the 0.5 km resolution dynamically downscaled surface winds with an average absolute direction difference of 43 and 20 for calm winds and moderate/strong winds respectively. Its average wind speed error is ~1.1 ms^{-1}. ~40 days are sufficient to create a representative database. Given the database, the procedure is extremely fast (a few seconds) and easy to operate, which makes it suitable for real-time non-expert fast-response applications.

1. Introduction

The atmospheric boundary layer flow is strongly affected by the terrain features. Therefore, when complex terrain is involved, high horizontal resolution is needed in order to properly describe the terrain features and to obtain the correct flow characteristics by numerical weather prediction (NWP) models. Two main approaches are available while calculating boundary layer flows: Dynamical downscaling and Statistical downscaling. Dynamical downscaling is well established and widely used (see e.g. Xue et al., 2014; Katsafados et al., 2011). It is a general method which can be applied anywhere over the globe. It starts the numerical calculation from a ~1–0.5° (~90–40 km) horizontal resolution data (boundary and initial conditions) provided by global circulation models (GCM). This data is processed while numerically solving the time dependent governing equations via nesting. A parent domain with 40–90 km horizontal resolution is defined. Afterwards the definition of sub-domains is made. Usually a 3:1 ratio is chosen while building the nested sub-domains, so that, each sub-domain covers smaller area and has a 3 times higher horizontal resolution relative to the previous domain. The innermost domain covers the area of interest and has the highest horizontal resolution. Along with increased horizontal resolution, comes the need to increase the time resolution, which is of the order of $\Delta x/U$, where Δx is the grid spacing and U is the characteristic advection velocity. The number of grid points in each grid and the size of the time step determine the amount of time and resources of these simulations. It turns out that the inner nests calculations consume the major part of the resources. This basic limitation poses a challenge whenever a real-time assessment is required.

In order to overcome the high load of dynamical downscaling, many statistical downscaling methods were derived (Sunyer et al., 2012; Lei et al., 2009; Wilby and Wigley, 1997; Wilby et al., 2004; Costa et al., 2008; Liu and Ren, 2014). These methods apply statistical relations between large scale atmospheric data and fine-resolution data in order to predict fine resolution meteorological variables. Analysis of a large set of past data from both scales reveals these relations and enables the prediction of the small scale data, given only the coarser resolution data. Statistical Downscaling algorithms use a variety of large scale data sources which varies from GCM to coarse resolution regional model predictions. The small scale data origin varies from fine-resolution NWP data to in-situ measurements. The sets of large scale variables used (predictors or explanatory variables) and small scale inferred variables (predictands or explained variables) depend on the application and differ accordingly. Often, the predictors are chosen according to some a-priori physical reasoning. The determination of the predictor variables also takes into account their spatio-temporal

E-mail address: alonm@iibr.gov.il (A. Manor).
[1] Tel.: +972 8 9381440; fax: +972 8 9381432.

http://dx.doi.org/10.1016/j.atmosres.2015.04.014
0169-8095/

range. The relations have to be derived and tested separately for each problem, on the basis of physical intuition, in a trial and error manner. However, this effort is worthwhile due to the significant saving in resources enabling real time applications. From this reason, we choose to develop a statistical downscaling algorithm for the prediction of surface winds. In this study our interest focused on complex terrain area, which dictates the use of high horizontal resolution (500 m and less) in order to get the correct prediction by dynamical downscaling.

Statistical downscaling for wind application has received considerable attention. Wind studies were developed mainly in two routes: 1. Long term forecasting relating to future climate due to global warming (Frey-Buness et al., 1995; Sailor et al., 2000; Wilby et al., 2004) 2. Short term forecasting, motivated by the rapid growth of wind energy use (see reviews by Foley et al, 2012; Lei et al, 2009; Costa et al, 2008).

Amongst the second group, many studies were dedicated to short-term wind speed forecasts. For example, Zhao et al. (2012) applied the WRF NWP model and ANN (Artificial Neural Network) algorithm to predict day ahead wind power. Salcedo-Sanz et al. (2009) applied the MM5 model and ANN algorithm to predict mean hourly wind speed. Nielsen et al. (2006) performed statistical downscaling based on principal components analysis with multiple regression. Kirchmeier et al. (2014) created surface wind-speed gamma distributions by linking large scale reanalysis data and in-situ measurements. Ishihara et al. (2006) and Yamaguchi and Ishihara (2003) created a downscaling algorithm to predict 10–20 m resolution surface winds by applying a CFD (Computational Fluid Dynamics) model.

The current study focuses on short-term wind vector forecasting. Fewer studies were dedicated to this aim. Salameh et al. (2009) statistically downscaled GCM data using linear regression, to yield surface wind vectors at a net of meteorological stations in a mountainous region in southern France. A combination of GCM and regional model data was used by Traveria et al. (2010) to generate a surface wind vector forecast at the Reus airport runway. The transformation from NWP data to local surface wind was done using an ANN method. Recently, Huang et al. (2014) applied linear regression to downscale coarse resolution reanalysis near-surface winds to a fine grid of 3 km spatial resolution.

This study conducted an ideal statistical downscaling experiment in order to develop a statistical downscaling algorithm for a real time prediction of surface winds over complex terrain. The Negev desert at southern Israel is a complex terrain desert with steep valleys and craters dictating the use of high horizontal resolution (300–500 m) in order to get the correct prediction of the surface wind flow. Many chemical plants are situated in this area and the need for a real time risk assessment tool is vital. To the best of our knowledge no statistical downscaling algorithm was developed for such a problem in this area. This work is a first step in this direction. It performs an ideal experiment in order to develop a statistical downscaling algorithm. Dynamical downscaling has been performed by the WRF model to create a historical database by the following two sets. The first set used 5 nested domains from 40.5 km to 0.5 km. The second set used 3 nested domains ranging from 40.5 km to 4.5 km.

The predictands were 0.5 km surface wind field derived from the first set. The second set was used to derive predictors from 4.5 km meteorological variables such as surface wind (speed and direction), temperature and surface fluxes. The statistical algorithm was built and verified by using the dynamically downscaled data. The skill of the proposed algorithm is tested by its ability to reproduce 0.5 km dynamically downscaled surface winds from 4.5 km dynamically downscaled data. The historical period included 180 representing days during 2009. The days were selected according to the abundance of synoptic states over the East Mediterranean area.

Unlike common analog algorithms applying minimal-differences methods in finding relevant analog, a Bayesian probabilistic approach, yields the most probable analog rather than the closest one. It is shown that the Bayesian algorithm improves surface wind reproduction.

Few applications use Bayesian inference to statistically downscale wind time series (Miranda et al. (2006), Jiang et al. (2013)), or perform wind field retrieval from satellite scatterometer observations (Cornford et al., 2004; Nabney et al., 2000). To the best of our knowledge, the Bayesian approach was not used for statistical downscaling of wind from NWP models data. The algorithm saves about 90% of the total runtime needed for the dynamical downscaling procedure and therefore, makes it suitable for real-time fast-response applications. Furthermore, running the algorithm is much simpler than running a typical NWP model. This enables usage by a non-expert.

The paper is structured as follows. Section 2 describes the datasets used in this work. Section 3 introduces the Bayesian aided analog statistical downscaling algorithm. Section 4 presents results and evaluations of the algorithm. A determination of the minimal database size for the Bayesian algorithm is performed. Section 5 concludes the study.

2. Data

2.1. The WRF regional atmospheric model: description and configuration

This work performs an ideal experiment in order to build a statistical downscaling algorithm. The database for this experiment was created by performing two separate sets of dynamical downscaling simulations. The simulations utilized the WRF model version 3.5 (Skamarock et al., 2008; http://www.wrf-model.org). The first set of runs used 5 two-way nested domains with horizontal resolutions of 40.5 km, 13.5 km, 4.5 km, 1.5 km and 0.5 km. The first domain covers the eastern coast of the Mediterranean Sea, and the innermost domain contains the northern Negev desert at the south of Israel (Fig. 1). The second set of runs used the first 3 two-way nested domains.

The WRF output from domain 5 defined the predictand data (0.5 km horizontal resolution). The predictor data, was derived from the third domain (4.5 km) meteorological fields of the second set.

27 eta levels were applied, 6 levels under 125 m above the ground. The model initial and boundary conditions were taken from the NCEP FNL (National Center for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce. 2000, updated daily. NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. http://dss.ucar.edu/datasets/ds083.2), having a time resolution of 6 h and horizontal resolution of 1°. The model was initialized at 0Z and run for 30 h during 180 selected days of 2009. Data was collected 6 h after initialization time, with 1 h time resolution.

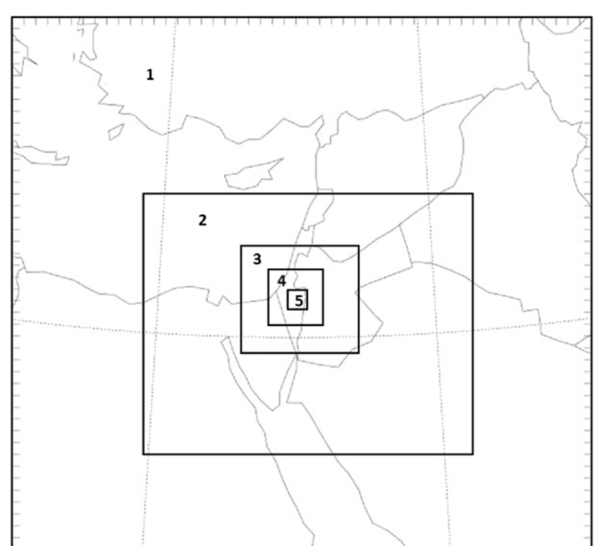

Fig. 1. The 5 nested domains used for the WRF simulations. Domain 1 covers the East Mediterranean area. Domain 5 is centered over the Negev desert at the south of Israel.