Forecasting of Surface Winds over Eastern Canada Using the Canadian Offline Land Surface Modeling System

LILY IOANNIDOU, WEI YU, AND STÉPHANE BÉLAIR

Environmental Numerical Prediction Research Section, Meteorological Research Division, Environment Canada, Dorval, Quebec, Canada

(Manuscript received 17 October 2012, in final form 24 February 2013)

ABSTRACT

The capability of the Canadian land surface external modeling system known as the Global Environmental Multiscale Surface (GEM-SURF) system with respect to surface wind predictions is evaluated. Based on the Interactions between Soil, Biosphere, and Atmosphere (ISBA) land surface scheme, and an exponential power law adjusted to the local stability conditions for the prediction of surface winds, the system allows decoupling of surface processes from those of the free atmosphere and enables high resolutions at the surface as dictated by the small-scale heterogeneities of the surface boundary. The simulations are driven by downscaled forecasts from the Regional Deterministic Prediction System, the 15-km Canadian regional operational modeling system. High-resolution, satellite-derived datasets of orography, vegetation, and soil cover are used to depict the surface boundary. The integration domains cover Canada’s eastern provinces at resolutions ranging from that of the driving model to resolutions similar to those of the geophysical datasets. The GEM-SURF predictions outperform those of the driving operational model. Reduction of the standard error and improvement of the model skill is seen as resolution increases, for all wind speeds. Further, the bias error is reduced in association with a rise in the corresponding value of the roughness length. For all examined resolutions GEM-SURF’s predictions are shown to be superior to those obtained through a simple statistical downscaling. In the prospect of the future development of a multicomponent system that provides wind forecasts at levels of wind energy generation, GEM-SURF’s potential for improved scores at the surface and its limited requirements in computer resources make it a suitable surface component of such a system.

1. Introduction

The importance of accurate wind prediction at the surface of the earth on time scales ranging from hours to days to weeks is an important aspect of weather prediction in view of the need for efficient and safe planning of human activities including transport, commerce, agriculture, and sport. It is also paramount for predicting accurate wind profiles in the lowest part of the atmospheric boundary layer, including the 50–100-m layer where wind turbines operate, and is therefore of interest to the wind energy generation sector. Surface winds are determined by an interplay of large-scale dynamics that govern circulations in the middle and upper atmosphere and surface boundary layer effects that are local in character. Among the latter are the orographic relief, thermal stability of the surface atmospheric layer, vegetation, and man-made obstacles in the vicinity of a site.

The forecasting of surface winds has been the focus of a number of modeling efforts. These typically involve limited-area models of high resolution in the horizontal, a large number of model levels in the lowest 1500 m in the vertical, and a more or less advanced assimilation scheme of surface observations (Roux et al. 2008; Lazić et al. 2010; Bernier and Bélair 2012; Wyszogrodzki et al. 2013). Identified areas for improvement in these modeling studies include the initialization of the land surface scheme of the model (Cheng and Steenburgh 2005) and the parameterizations in the model’s boundary layer (Santos-Alamillos et al. 2013). Sensitivity to the terrain morphology was identified as a prime source of error.

Corresponding author address: Lily Ioannidou, Environmental Numerical Prediction Research Section, Meteorological Research Division, Environment Canada, 2121 Trans-Canada, Dorval QC H9P 1J3 Canada.

E-mail: lilyioan@yahoo.com

DOI: 10.1175/JAMC-D-12-0284.1
(Zhang et al. 2013), and a number of studies (Mass and Ovens 2010; Jiménez and Dudhia 2012) focus on improvements in the representation of the resolved and in the parameterization of the unresolved topographic effects in regions of complex terrain. Sensitivity of error was not solely identified in relation to terrain variability but to geographical variations in general (Wyszogrodzki et al. 2013), as well as to the season and time of day. The large sensitivity to the local conditions identified in the aforementioned studies suggests that for an accurate representation of the boundary layer land–atmosphere exchanges of heat, moisture, and momentum that determine the distribution of the surface wind a high-resolution representation of small-scale variations of the surface features including orography, vegetation cover, soil type, snow and ice cover, coastline, and atmospheric stability is imperative.

In recent years, an external land surface modeling system [the Global Environmental Multiscale Surface (GEM-SURF) system] has been developed at the Meteorological Service Canada of Environment Canada (EC). The main objective of the system is to improve Environment Canada’s operational prediction of surface variables in a cost effective way. To achieve the high resolution needed to accurately simulate land surface processes, the system operates on two levels. First, the output of an operational three-dimensional atmospheric model is used to extract the low-level atmospheric fields that will provide the regular forcing for the land surface modeling system, and a high-resolution downscaled version of these forcing fields is generated. The land surface component of GEM-SURF, driven by the downscaled fields, is integrated in a second phase using a grid as small as necessary for a realistic representation of the small-scale heterogeneities of the surface, including orographic, vegetation, land–sea contrast, and soil type surface variations. As a consequence of the representation of the surface geophysical fields on a fine grid, a finescale and more accurate representation of the distribution of the surface fluxes of heat, humidity, and momentum is expected by the GEM-SURF integration than that predicted by the operational model that is driving it.

It is hypothesized in this study that by introducing finescale information at the surface and thus reproducing the fine scales that characterize the surface–atmosphere exchanges, the land surface modeling system is able to improve on the accuracy of surface wind predictions relative to a lower-resolution weather prediction model. The range of resolutions appropriate for the offline simulations is determined by the resolution of the available geophysical datasets representing orography, vegetation cover, land–sea distribution, and so on but also by the density of the validating observational network and the constraints imposed by the model’s surface parameterizations. An obvious advantage of this approach in relation to three-dimensional atmospheric models operating in the same range of high resolutions is its significantly lower computational cost.

In earlier studies (Carrera et al. 2010; Bernier et al. 2011; Leroyer et al. 2011), the GEM-SURF system’s performance with respect to the prediction of screen-level temperature, terrestrial snow, and surface fluxes in an urban setting was investigated and shown to be improved compared with Canadian operational models. In the prospect of developing a multicomponent modeling system that provides reliable forecasts of wind profiles in the surface boundary layer in regions with wind energy generation potential, at a low computational cost, the objective of this study is to investigate GEM-SURF’s capabilities with respect to surface wind predictions. More particularly, to evaluate the impact that increasing GEM-SURF’s resolution could have in obtaining improved surface wind prediction scores in comparison 1) with models currently operational at the Canadian Meteorological Centre and 2) with a downscaling obtained through interpolation and bias-based corrections of the operational predictions. The domain covers Canada’s eastern provinces and the period of the simulation extends over the 6-month period from 1 September 2008 to 28 February 2009. The results presented in this study refer to the first three months of the simulation from 1 September 2008 to 30 November 2008. These correspond to the autumn period, before processes in the land surface scheme of GEM-SURF that involve changes of the phase of the water present in the soil/vegetation/snow, become active.

The paper is organized as follows: In section 2 the GEM-SURF system is presented with emphasis on the surface wind prediction component, and also the setup of the experiment and the dataset used for evaluation of the results. In section 3 the mean surface wind forecast scores are presented. The dependence of the scores on the wind magnitude and the sensitivity to increasing resolution are discussed. Conclusions are presented in section 4.

2. Model simulations

a. The GEM-SURF model

GEM-SURF is an external land surface modeling system in which surface variables such as surface temperature and other characteristics evolve in a prognostic manner. Using low-resolution atmospheric forcing that is first interpolated on a high-resolution grid and then adjusted vertically based on high-resolution orographic
and physiographic information, GEM-SURF produces forecasts of the surface variables on a kilometer and subkilometer scale.

The land surface modeling component in GEM-SURF is based on the Interactions between Soil, Biosphere, and Atmosphere (ISBA) land surface scheme (Noilhan and Planton 1989; Noilhan and Mahfouf 1996; Boone et al. 1999, 2000). The version of ISBA implemented in the Canadian operational models (Bélaire et al. 2003a, b) generates predictions for 10 surface variables. These include the surface temperature and soil temperature, the soil volumetric water and ice contents, the water retained on the vegetation canopy and in the snowpack, the snow mass, the snow depth and snow density, and the albedo of the snow. GEM-SURF also predicts the turbulent fluxes of heat and momentum in the surface atmospheric layer following classical aerodynamic equations (Delage et Girard 1992; Delage 1997). Based on the predictions for the fluxes, stability functions $\Phi_h$ and $\Phi_m$ of heat and momentum are calculated that adapt the predicted near-surface wind to the temporally and spatially varying conditions of thermal and mechanical stability (Noilhan and Mahfouf 1996).

The formulation of $\Phi_h$ and $\Phi_m$ is based on local values of the Richardson number and of the roughness length parameters $Ri$ and $z_0$, so that a strong sensitivity of the predicted surface wind is ensured to diurnal or seasonal variations in the stability of the near-surface atmospheric layer and to regional variations in the roughness of the underlying surface. The dependence on $Ri$ and $z_0$ is such that 1) the predicted surface wind drops for large $Ri$ (thermally stable) and rises for small or negative $Ri$ (thermally unstable layer) and 2) increases for small $z_0$ (smooth terrain) and diminishes for large $z_0$ (rough terrain).

A basic limitation in the current formulation of $\Phi_h$ and $\Phi_m$ is associated with the absence of coupling in GEM-SURF of the near-surface processes with the free atmosphere. In the absence of a two-way exchange of heat and momentum that would establish a free equilibrium between the near-surface atmospheric layer and the layers above, the turbulent fluxes of heat and momentum are subject to ad hoc enhancement/weakening to maintain the surface temperatures at realistic values at different times of the day. The need to apply corrections to the surface fluxes is a common issue in offline land surface models (Schulz et al. 1998; Abramowitz et al. 2007). Although the effect of the corrections is beneficial in that it reduces the cold bias that typically characterizes the surface temperatures (De Haan et al. 2007), it poses constraints on the predicted surface winds. These are discussed in the interpretation of the results of the GEM-SURF integrations.

b. Experimental design

1) COMPUTATIONAL DOMAIN

Three large domains of different size and resolution (Fig. 1a) were designed. These cover eastern Canada, namely the provinces of Quebec, New Brunswick, Nova Scotia, Prince Edward Island, and, in large part, Newfoundland. A sector of the Atlantic Ocean is included in the largest of the three domains. From largest to smallest the grid spacing is 15, 2.5, and 1.0 km. Each domain was designed to be smaller and nested within the domain of the immediately lower resolution. The three domains offer the opportunity of testing GEM-SURF’s land surface modeling schemes under a wide range of physiographic conditions. In the smallest of the three domains where the observing stations are situated, elevations range from sea level to 1300 m, and the vegetation cover includes a number of different types, namely cropland, grassland, shrubs, swamps, deciduous trees, evergreen forests, and other. With respect to the land–water distribution, a characteristic feature in the domains’ geography is the rich coastline and the ubiquitous presence of inland water bodies.

In the prospect of GEM-SURF forming part of a multicomponent near-surface modeling system appropriate for wind energy predictions, three smaller domains were also designed and placed selectively over the sites of wind farm projects. These projects have become progressively operational in the province of Quebec between 2006 and 2008 with a total number of 215 turbines and a total generation capacity of 320 MW. The selected sites are of Anse-à-Valleau, Carleton, and Baie-des-Sables and the respective domains, delineated in Fig. 1a by the turquoise, magenta, and green rectangles, have sizes of approximately 12 500, 8400, and 6100 km$^2$. The grid spacing for the three small domains is 250 m.

2) SURFACE GEOPHYSICAL FIELDS

To specify values at each grid cell of the respective domain for the orographic and vegetation component of the roughness length parameter, and for the input coefficients to ISBA like the thermal conductivity, leaf area index, and the coefficients of the water retention curves, a set of satellite-derived geophysical databases was used to provide an accurate depiction of the fields of elevation, land–water mask, vegetation, and soil type. The criterion in the selection of the fields is that they combine high resolution with maximum reliability, and to this purpose the set includes recently acquired high-resolution databases as well as older lower-resolution databases still used in various operational models.

A concise view of the databases is presented in Table 1. As seen in Table 1, for the elevation, the primary database
FIG. 1. (a) The six domains for the GEM-SURF integrations. For the three large domains, grid spacings are 15, 2.5, and 1.0 km. The grid spacing for the three small domains centered over Anse-à-Valleau (turquoise rectangle), Carleton (magenta), and Baie-des-Sables (green) is 250 m. (b) Sites of the 72 EC observing stations (blue and red crossed circles) used for the evaluation of the 15-, 2.5-, and 1.0-km integrations, superimposed on the orography (m) of the 2.5-km domain. Blue circles indicate the subsample of 10 stations used for the evaluation of the 250-m experiments; red circles indicate the remaining 62 stations. The color scale represents elevation contour intervals of 100 m.
that was used is the Canadian Digital Elevation Data 1:250,000 (CDED250) dataset. This database was compiled by Natural Resources Canada and provides elevation data over Canada at an average resolution of 50 m in the east–west direction and 90 m in the north–south direction. Over non-Canadian territory and over the CDED250 void regions of Canada, CDED250 was complemented by the Shuttle Radar Topography Mission (SRTM) digital elevation dataset and the U.S. Geological Survey (USGS) seamless elevation dataset, at 90- and 900-m resolutions, respectively. For the vegetation type and cover, the USGS-Global Land Cover Characteristics (GLCC) is used, also at 900-m resolution. The USGS-GLCC defines the vegetation cover over the North American continent in terms of 202 land-use/land-cover classes that are subsequently remapped onto 26 vegetation classes for ISBA. The mapping of different land-use/vegetation coverage types onto the 26 classes, shown in Table 2 along with the corresponding values of the roughness length, is a joint product of the Canadian Meteorological Center and Météo-France, compiled from a number of earlier land-use classification tables that were tested and used in the two centers. Of the remaining geophysical datasets, the one describing the soil type coverage is obtained, as shown in Table 1, from the Food and Agriculture Organization of the United Nations (FAO) database at 1° resolution, and the dataset describing the land–water mask is obtained from the GlobCover database at 300-m resolution. The above land–water mask dataset was then subjected to a number of checks to ensure its consistency with the land–water mask fields given by CDED250 and USGS-GLCC.

The aforementioned datasets are first subjected to a sequence of consistency checks to identify mutual inconsistencies particularly with respect to the land–water mask distribution. A remapping is then performed from the grid of the geophysical field onto the GEM-SURF grid. The remapping consists of identifying the lines of intersection between the grid of the geophysical dataset and that of GEM-SURF and computing an average over the data points, or grid cells, of the geophysical dataset that are contained within a given GEM-SURF grid box. When the difference between the GEM-SURF resolution and that of the geophysical dataset is large, an error is expected in association with the remapping procedure, as the computed average point value cannot accurately represent the geophysical field's variability within the grid box. In these cases, alternative methods of remapping in which a measure of the variability within the grid box is included in the point estimate

<table>
<thead>
<tr>
<th>Geophysical datasets</th>
<th>Primary</th>
<th>Complementary 1</th>
<th>Complementary 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elevation</td>
<td>CDED250</td>
<td>SRTM</td>
<td>USGS</td>
</tr>
<tr>
<td></td>
<td>50 m east–west and 90 m north–south</td>
<td>90 m</td>
<td>900 m</td>
</tr>
<tr>
<td>Land–water mask</td>
<td>GlobCover</td>
<td>CDED250</td>
<td>USGS-GLCC</td>
</tr>
<tr>
<td></td>
<td>300 m</td>
<td>90 m</td>
<td>900 m</td>
</tr>
<tr>
<td>Vegetation type</td>
<td>USGS-GLCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>900 m</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soil type</td>
<td>FAO</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1°</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 1.** The geophysical datasets, and their respective resolutions, used to define the surface fields of elevation, land–water mask, vegetation, and soil type for the GEM-SURF integrations. The first data column shows the dataset used as the primary source; columns 2 and 3 show the complementary datasets.

<table>
<thead>
<tr>
<th>Class</th>
<th>Vegetation type</th>
<th>$Z_0$ (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Water</td>
<td>0.0001</td>
</tr>
<tr>
<td>2</td>
<td>Ice</td>
<td>0.0001</td>
</tr>
<tr>
<td>3</td>
<td>Inland lake</td>
<td>0.0001</td>
</tr>
<tr>
<td>4</td>
<td>Evergreen needleleaf trees</td>
<td>1.5</td>
</tr>
<tr>
<td>5</td>
<td>Evergreen broadleaf trees</td>
<td>3.5</td>
</tr>
<tr>
<td>6</td>
<td>Deciduous needleleaf trees</td>
<td>1.0</td>
</tr>
<tr>
<td>7</td>
<td>Deciduous broadleaf trees</td>
<td>2.0</td>
</tr>
<tr>
<td>8</td>
<td>Tropical broadleaf trees</td>
<td>3.0</td>
</tr>
<tr>
<td>9</td>
<td>Drought deciduous trees</td>
<td>0.8</td>
</tr>
<tr>
<td>10</td>
<td>Evergreen broadleaf shrub</td>
<td>0.05</td>
</tr>
<tr>
<td>11</td>
<td>Deciduous shrubs</td>
<td>0.15</td>
</tr>
<tr>
<td>12</td>
<td>Thorn shrubs</td>
<td>0.15</td>
</tr>
<tr>
<td>13</td>
<td>Short grass and forbs</td>
<td>0.02</td>
</tr>
<tr>
<td>14</td>
<td>Long grass</td>
<td>0.08</td>
</tr>
<tr>
<td>15</td>
<td>Crops</td>
<td>0.08</td>
</tr>
<tr>
<td>16</td>
<td>Rice</td>
<td>0.08</td>
</tr>
<tr>
<td>17</td>
<td>Sugar</td>
<td>0.35</td>
</tr>
<tr>
<td>18</td>
<td>Maize</td>
<td>0.25</td>
</tr>
<tr>
<td>19</td>
<td>Cotton</td>
<td>0.1</td>
</tr>
<tr>
<td>20</td>
<td>Irrigated crops</td>
<td>0.08</td>
</tr>
<tr>
<td>21</td>
<td>Urban</td>
<td>1.35</td>
</tr>
<tr>
<td>22</td>
<td>Tundra</td>
<td>0.01</td>
</tr>
<tr>
<td>23</td>
<td>Swamp</td>
<td>0.05</td>
</tr>
<tr>
<td>24</td>
<td>Desert</td>
<td>0.05</td>
</tr>
<tr>
<td>25</td>
<td>Mixed shrubs</td>
<td>1.5</td>
</tr>
<tr>
<td>26</td>
<td>Mixed-wood forest</td>
<td>0.05</td>
</tr>
</tbody>
</table>

**Table 2.** Vegetation/land-cover classes used by ISBA and corresponding roughness lengths $Z_0$. 
would be beneficial in reducing the error. As model resolution increases and becomes comparable to that of the geophysical dataset(s), the above error in representativity becomes smaller.

3) ATMOSPHERIC FORCING

The atmospheric fields needed as input for the land surface model, referred to as the forcing fields, are obtained from operational forecasts of the Canadian Regional Deterministic Prediction System (RDPS; Mailhot et al. 2006), based on the GEM model. The RDPS provides 48-h forecasts over a domain covering North America and part of the adjacent oceans with a 15-km grid spacing. The system uses a terrain-following vertical coordinate η. The forcing fields include information on incident radiation, hourly precipitation, and surface pressure as well as air temperature, wind, and specific humidity at the η level closest to the surface, where η = 0.995. For the period covered in this study extending from 1 September 2008 to 30 November 2008, the RDPS integrations and the surface variables’ fields are initialized at 0000 and 1200 UTC daily. To avoid spinup errors and errors due to loss of skill with increased lead time, only forecasts valid from 6 h up to 17 h after initialization were used.

Before being introduced into GEM-SURF, the forcing fields are submitted to a preprocessing that involves horizontal interpolation onto the GEM-SURF computational grid. The interpolation is cubic for the temperature, pressure, humidity, and wind fields and linear for the precipitation rate and the incident radiation. It is followed for the temperature, pressure, humidity, and precipitation by an “elevation based” adjustment, an adjustment shown (Dodson and Marks 1997; Stahl et al. 2006; Sheridan et al. 2010; Bernier et al. 2011) to be especially advantageous in mountainous regions where discrepancies between the smooth orography depicted by a low-resolution grid and the rugged orography seen by a high-resolution grid are largest.

The adjustment procedure consists of applying a height-based correction to the value of the field at each grid point so that the difference in elevation of the point on the 15-km-resolution RDPS grid and the higher-resolution GEM-SURF grid is considered. For the temperature, the correction involves a mean cooling/warming of the forcing temperature when the difference between a point’s altitude on the high-resolution grid and its altitude on the low-resolution grid is positive/negative. A mean lapse rate of 6.0 K km⁻¹ was used following the work of Bernier et al. (2011), where the 6.0 K km⁻¹ value was found to represent the best linear fit to the observed relation between screen-level temperature differences and elevation differences estimated over all possible pairs of hourly observations in the November 2007 to May 2009 period, at the 51 sites that constituted the Vancouver Winter Olympics observing network. Given the large range of elevations (ranging from sea level to 1650 m) covered in Bernier et al. (2011), the above mean value for the lapse rate was considered also suitable for the purpose of this study that aims at evaluating model performance in the mean. It was considered that fine-tuning of the lapse rate to location, season, and hour of day that is required for large domains like in Fig. 1a, characterized by diversity of environmental conditions, forms the objective of an independent study and was not undertaken.

Following the temperature, the surface pressure is adjusted so that the “new” adjusted temperature and the “new” adjusted surface pressure are in hydrostatic balance. The specific humidity is adjusted so that the relative humidity is conserved for the adjusted temperature and surface pressure, and the precipitation is adjusted as a function of the new temperature, that is, to snow/rain if the adjusted temperature is below/above 0°C. The incident radiation is not adjusted. The forcing winds are not explicitly submitted to adjustment; however, they are partly, via the stability functions Φh and Φm, subject to the effects of the adjusted temperature that enters into the calculation of Φh and Φm.

4) MODEL INTEGRATIONS

The GEM-SURF integrations were initialized at 0000 UTC 20 July 2008 allowing for a 40-day spinup. The time step was set to 30 min. The forcing RDPS fields valid at 0000, 0600, 1200, and 1800 UTC were linearly interpolated in time to provide hourly input. For the sea surface temperature and the continuous sea ice fraction, climatological fields were used and updated at 0000 UTC daily. The choice of climatological fields for the SST and sea ice was dictated by the insensitivity of the GEM-SURF predictions to the details of the sea surface temperature and sea ice representations, given the model’s one-dimensional character. At the surface, the 10 prognostic fields of ISBA were specified at the initial time of the experiment from the surface analysis of RDPS at that time. The latter is based on a sequential assimilation scheme (Bélair et al. 2003a) using screen-level observations of temperature and humidity. From the initial time onward, the GEM-SURF simulations proceed in a continuous “open loop” mode with no further assimilation of surface variables. A possibility of model drifting as the integration proceeds in the absence of any observational input at the surface exists; however, this is small since the RDPS-based forcing fields are subject to data assimilation.
Sample surface wind forecasts from the integrations with 15-, 2.5-, and 1.0-km grid spacing are shown in Figs. 2a–c. The forecasts are valid for 1400 eastern standard time (EST) 28 November 2008. The domain shown is centered over the Gaspé Peninsula, located in the eastern part of the province of Quebec. The predicted surface wind speed is shown in color as a scalar quantity and by the arrows as a vector. The impact of the additional details in the representation of the geophysical fields as the grid spacing decreases from 15 to 1.0 km on the finescale structure of the surface wind field can be seen over land. The difference is more pronounced between the 15- and 2.5-km scalar wind speed distributions, but differences are also discernible for the small change from 2.5- to 1.0-km grid spacing. Over the water masses surrounding the peninsula, where the geophysical fields have no impact and the forecasts are subject only to the effect of horizontal interpolation, it can be seen that the forecasts from the 15-, 2.5-, and 1.0-km integrations are quite similar.

3. Observational data and model performance

a. Data for objective evaluation

The results of the GEM-SURF simulations were evaluated against single point wind measurements made in the period from 1 September 2008 to 30 November 2008 at observing stations of Environment Canada located in the provinces of Quebec, New Brunswick, Nova Scotia, Prince Edward Island, and Newfoundland. A cubic interpolation was made of the surface wind forecasts from the surrounding grid points on the site of each observing station for every hour of the 3-month period. For the experiments at 15, 2.5, and 1.0 km, the 72 EC observing stations lying within the 1.0-km domain were used (Fig. 1b). For the 250-m experiment, the 10 EC stations shown by the blue dots were used.

The station measurements consist of hourly recordings of the observed surface wind. They represent quasi-instantaneous values, that is, averages of continuous recordings taken over the 2 min surrounding the hour. For some stations, the recordings are made at 20, 30, or 40 min past the hour. In these cases they are linearly interpolated onto the hour. Quality checks were made for identifying at each station’s record subperiods for which the measurements are incomplete or unreliable. These were excluded from the evaluation.

b. Scores

A first evaluation of the GEM-SURF predictions is made by comparing the frequency distributions of forecast and observed wind speeds for the observed range of winds. Figure 3a shows the frequency distributions of observed (gray) and forecast winds from the 15-km- (blue), 2.5-km- (green), and 1.0-km- (red) resolution integrations. The curves in Fig. 3b represent the corresponding fifth-order polynomial approximations obtained with regression coefficient $R^2$ values close to 0.996. The curves are characterized by positive skewness with their frequency maximum at low speeds and extensive tail at high wind speeds. Comparison between observed and forecast frequencies in Fig. 3a in the $0 \leq U \leq 4 \text{ m s}^{-1}$ range indicates a shift of the frequency maximum to higher speeds in the forecasts relative to the observations. Specifically, the frequency maximum is displaced from speed 2 m s$^{-1}$ in the observations to 3 m s$^{-1}$ in the forecasts. This is not surprising given the general tendency of models to overestimate the lighter winds. In the corresponding polynomial curves (Fig. 3b) the shift of the low speed frequency maximum to the right is less pronounced but still evident. For speeds $U \geq 5 \text{ m s}^{-1}$, forecast and observed frequencies are in good agreement in both the histogram and the polynomial representation, and it can be seen that the agreement between the two improves as resolution increases. The 1.0-km-resolution curve in particular (Fig. 3b) closely follows for speeds $U \geq 5 \text{ m s}^{-1}$ the frequency curve of the observed winds. For the very low winds $0 < U < 2 \text{ m s}^{-1}$, Fig. 3b indicates that the curves that correspond to the 1.0- and 2.5-km-resolution forecasts lie closer to the observed frequency curve, in comparison to the curve of the 15-km-resolution forecast. Departures between observed and forecast distributions are larger in the intermediate range of $2 \leq U \leq 4 \text{ m s}^{-1}$.

To further analyze the error distribution in each bin of the observed wind speed range, conditional quantile plots for the 15- and 1.0-km-resolution forecasts are shown in Figs. 3c,d. For both resolutions it is seen that the spread between quantiles is within acceptable limits and its variation with wind speed is small. For example, for a forecast wind speed $U = 5 \text{ m s}^{-1}$ in the 1.0-km-resolution integration (Fig. 3d), the interquartile range of observed winds is from 4.5 to 6.0 m s$^{-1}$, while for forecast wind speed $U = 10 \text{ m s}^{-1}$ the interquartile range is 8.5–12.5 m s$^{-1}$. An estimate of an upper limit for the magnitude of the error that is associated with the above interquartile range can be obtained by considering the case of observed wind of $6 \text{ m s}^{-1}$ for the forecast wind of $5 \text{ m s}^{-1}$ generated from the 1.0-km integration. The departure of $1 \text{ m s}^{-1}$ between the two represents a relative error of 15%–20% of the observed wind value. In the case of a forecast wind of $U = 10 \text{ m s}^{-1}$, and an observed wind of $12.5 \text{ m s}^{-1}$ that corresponds to the upper limit of the respective interquartile range, the departure of $2.5 \text{ m s}^{-1}$ also represents a relative error of
15%–20% of the observed wind value. The latter is similar to the upper limit of the relative error estimated for the forecast wind of 5 m s\(^{-1}\). As for the interquartile range, estimates from Figs. 3c and 3d for the upper limits of errors associated with the interdecile range indicate very little sensitivity of forecast error to the wind speed magnitude.

Comparison between Figs. 3c and 3d indicates a superiority of the 1.0-km-resolution forecast over the entire range of observed winds. First, it is noted that the interquartile and interdecile ranges of the 1.0-km-resolution integration are more symmetric about the “perfect” forecast than the respective ranges of the 15-km-resolution integration. The 0.5-quantile curve crosses the perfect forecast at wind speed values equal to 5.8 and 3.3 m s\(^{-1}\) for the 15- and 1.0-km integrations, indicating that the respective forecasts overestimate/underestimate the observed winds for speeds smaller/larger than the above value(s). Of the two values, 3.3 indicates smaller departures between forecast and observed frequencies in the low wind speed range \(0 \leq U \leq 4\) m s\(^{-1}\) and is consistent with the better approximation of the observed frequency curve by the 1.0-km forecast curve in the above range seen earlier. Most important, it is seen that for the entire range of observed wind speeds the departures of the 0.5-quantile curve from the perfect forecast are smaller for the 1.0-km-resolution forecast than for the 15-km-resolution forecast.

For further evaluation of the GEM-SURF predictions the driving atmospheric model RDPS operating at 15-km resolution is used as reference for assessing the impact of increasing the resolution in the integrations. Mean surface wind forecast scores are calculated, averaged over the 72 observing sites for the 3 months of the integration period. The metrics presented are the root-mean-square error (RMSE) of the forecast surface winds, the standard deviation of the error (STDE), and the error bias.

An overview of the geographical distribution of the error within the domain of the experiment is given first, in Fig. 4. The figure shows the mean daily RMSE of the predicted surface winds on the EC observing sites, averaged over the period of the integration. The RMSE values correspond to the 15-km-resolution integration. It is seen that values range between 1.2 and 2.4 m s\(^{-1}\). Lower values are seen in the central part of the domain, a region characterized by low elevations; larger values are seen in the northern and southern parts, regions characterized by higher elevations and steep gradients in proximity to the coastline.

Next, the mean diurnal evolution of surface wind is shown as predicted by the 15-, 2.5-, and 1.0-km grid GEM-SURF integrations (Fig. 5a). The observations show surface winds around 4.0 m s\(^{-1}\) in the evening/nighttime,
rising to a peak value of 5.0 m s\(^{-1}\) in the warmest hours of day. Increase in resolution in the examined range of 15–1.0 km leads to decrease in the predicted wind speeds, consistently at all the hours of day. Also evident in Fig. 5a is the underestimation by the model of the amplitude of the observed diurnal cycle.

Examination of the corresponding curves of the mean diurnal evolution of the RMSE, shown by the solid lines in Fig. 5b, show a daily mean RMSE of 2.25 m s\(^{-1}\) for the 15-km integration. This is reduced to 1.98 m s\(^{-1}\) in the 2.5-km integration. A further reduction of 0.1 m s\(^{-1}\) is seen for grid spacing of 1.0 km. The shaded zones measure the spread around each RMSE curve obtained with the bootstrap method and a trial number of 1000. The confidence interval was set to 95%, giving limits of the bootstrap distribution between the 2.5th and the 97.5th percentiles. The bootstrapping was applied to the 72 stations’ forecast time series, each time series consisting of wind forecasts at the same hour of day for all days of the sample. It is seen that the shaded zones surrounding the 15- and 2.5-km curves do not overlap, suggesting that the reduction in RMSE from 15 to 2.5 km is significant. The 2.5- and 1.0-km shaded zones overlap, especially in the warm part of the day, but the facts that the increment in resolution from 2.5 to 1.0 km is small and that the overlapping during hours other than the midday is small suggest that the reduction in RMSE with resolution is significant for the examined resolution range from 15 to 1.0 km.

Comparison with the operational predictions by the RDS running at 15-km resolution is shown in Fig. 5c. The metric shown is the mean-square error skill score (MSESS) defined as

\[
\text{MSESS} = 1 - \left( \frac{\text{MSE}_{\text{GEM_SURF}}}{\text{MSE}_{\text{RDPSS}}} \right),
\]

where MSE\(_{\text{GEM_SURF}}\) is the mean-square error of the GEM_SURF predictions, and MSE\(_{\text{RDPSS}}\) is the mean-square error of the RDS predictions, used as reference for the evaluation. It is seen that the blue curve (15-km-resolution forecast) vacillates around 0.0, indicating
comparable skill of the 15-km GEM-SURF to the 15-km RDPS run. Small differences in performance between the two forecasts stem from the use of different sets of geophysical fields and from the assimilation of near-surface variables in the RDPS integration. As resolution rises to 2.5 and 1.0 km, the MSESS scores rise to 0.07 and 0.11, respectively, consistent with the reduction in the RMSE in Fig. 5b. It is concluded that by using higher resolutions in the GEM-SURF integrations, a gain in the near-surface wind forecast skill is achieved relative to the operational predictions.

It is considered that the reduction in the RMSE as resolution increases is associated with the increased accuracy in the representation of the physiographic information at the surface. The latter leads to more accurate values for the GEM-SURF parameters associated with orography, vegetation, and soil that lead to more realistic distribution of surface fluxes and hence to more accurate surface wind predictions. The increased accuracy in the orography representation in particular has an additional impact as it improves the quality of the elevation adjustment of the forcing fields.

The reduction in error with increasing resolution is also seen when only the random errors’ contribution to the total error is considered by removing the bias component. The diurnal curves of the standard deviation of the wind error, shown by the dashed lines in Fig. 5b, indicate an improvement in forecast skill with increasing resolution of 0.1 m s$^{-1}$ when going from the 15-km to the 2.5-km run, and another 0.05 m s$^{-1}$ from 2.5 to 1 km. The contribution of the bias to the total RMSE is small in general and becomes smaller during daylight hours as seen by the RMSE and STDE curves that almost coincide at midday, suggesting that almost all the RMSE is associated with random errors.

Consistent with the STDE, the corresponding curves for the wind bias (Fig. 5d) show a constant positive bias in the evening/nighttime, reflecting the wind overestimates during those hours (Fig. 5a), and a prominent minimum in the form of an extensive dip during daylight. The decrease of the daily mean bias value by a constant amount as resolution increases from 15 to 1.0 km suggests for each increment in resolution a negative wind speed offset. Investigation of the bias behavior revealed a relation between the change with resolution of the mean daily bias value and the corresponding change in the value of the roughness length $Z_0$. The latter is shown in Table 3. The rise in $Z_0$ as resolution increases is consistent with the decrease in the mean forecast wind speed (Fig. 5a), as a larger $Z_0$ implies a stronger drag on the predicted surface wind.

A comparison is made in Fig. 5e between the improvement in the GEM-SURF predictions with increasing resolution and the improvement obtained through a statistical downscaling of the operational RDPS 15-km forecast. The comparison is motivated by the need to demonstrate the value of the offline land surface system integration with regard to the contribution of the physical processing of the surface variables by the model schemes to the quality of the forecast. The objective is to show that the model-generated wind forecast at a given high resolution is superior to a forecast based on statistical downscaling of surface winds predicted by the operational RDPS 15-km forecast to that same resolution. This comparison is especially needed since the forcing fields for the high-resolution GEM-SURF runs are obtained through downscaling of the RDPS 15-km winds at the forcing level.

The downscaling consists of cubically interpolating the RDPS 15-km surface wind forecast fields on the 2.5- and 1.0-km grids and subtracting at each gridpoint the mean bias value for the respective resolution and hour of day as obtained from Fig. 5d. Figure 5e shows the RMSE (solid lines) and STDE (dotted) for the GEM-SURF predictions at 2.5 km (green) and 1.0 km (red) resolutions together with the RMSE and STDE of the forecasts obtained by downscaling to 2.5 km (brown) and to 1.0 km (light green). It is evident that the GEM-SURF predictions outperform those obtained by the downscaling. For the 2.5-km predictions the GEM-SURF RMSE is on average 0.1 m s$^{-1}$ smaller than that of the downscaled field, while for the 1.0-km predictions it is 0.17 m s$^{-1}$ smaller. It is concluded that the improvement attained by GEM-SURF as resolution increases accounts for more than just application of a bias correction on
a downscaled low-resolution surface forecast. The latter in turn implies that the physical treatment of the surface and near-surface quantities by the model operating at high resolution is a significant component of the observed improvement in skill with increasing resolution.

In view of the dependence of the surface wind predictions on the stability functions of heat and momentum $\Phi_h$ and $\Phi_m$, and of the direct impact of the latter on the surface temperature forecasts, the characteristics of the GEM-SURF surface wind forecasts are compared in Fig. 5f to the corresponding features of the temperature forecasts. The surface temperature prediction curves are more satisfactory than the winds as 1) the amplitude of the observed diurnal temperature variation is well simulated and 2) the forecasts are relatively insensitive to the rise in resolution. The observed discrepancy between wind and temperature forecasts is attributable to the weight given in the present form of $\Phi_h$ and $\Phi_m$ (Delage and Girard 1992) to the maintenance of realistic surface temperatures. In the absence of two-way exchanges of heat and momentum between the surface and the upper tropospheric layers, the nighttime/daytime turbulent exchanges obtained through $\Phi_h$ and $\Phi_m$ in the

![Fig. 5](https://example.com/figure5.png)

**Fig. 5.** (a) Average diurnal evolution of the surface wind predicted by GEM-SURF (color) and of the observed surface winds (black) for the 3 months of the integration period. Results from the 15-, 2.5-, and 1.0-km runs are in blue, green, and red, respectively. The curves represent averages over the 72 Environment Canada observing stations. The abscissa represents local hour of day from 2000 of one day to 1900 of the following. The ordinate represents wind speed in meters per second. (b) Average diurnal evolution of the RMSE (solid lines) and the STDE (dashed) of the surface winds predicted by GEM-SURF for the 3-month integration period; the shaded zones around each curve represent error margins for a 95% confidence interval, obtained with the bootstrap method with a trial number of 1000; color convention is as in (a). (c) Average diurnal evolution of the autumn MESS of the GEM-SURF 15- (blue), 2.5- (green), and 1.0-km (red) integrations using as reference, the operational predictions of the 15-km-resolution RDPS integrations. (d) As in (b), but for the wind error bias. (e) Diurnal evolution of the RMSE (solid) and STDE (dashed) of the GEM-SURF 2.5- (green) and 1.0-km (red) predictions, and of the RDPS 15-km predictions, interpolated on the 2.5- (light green) and 1.0-km (brown) grids and corrected for the respective mean bias. (f) As in (a), but for the forecast and observed surface temperatures; units on the ordinate are degrees Celsius.

### Table 3. Average values of the roughness length $Z_0$ (m) for each of the examined resolutions for the sample of the 72 observing stations lying within the large domains and for the 10 observing stations situated within the three smaller domains.

<table>
<thead>
<tr>
<th>Resolution</th>
<th>stations</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 km</td>
<td>0.445</td>
</tr>
<tr>
<td>2.5 km</td>
<td>1.175</td>
</tr>
<tr>
<td>1.0 km</td>
<td>1.84</td>
</tr>
<tr>
<td>250 m</td>
<td>1.94</td>
</tr>
</tbody>
</table>
surface model, are too strong/weak, resulting in quite realistic temperatures but over/underestimates of the surface winds at the respective times. It is considered that a reformulation of $\Phi_h$ and $\Phi_m$, so as to improve on the wind forecasts without deteriorating the quality of the surface temperature forecasts, is required. A more elaborate treatment of the stability functions, within a multilayer boundary layer model, is among the aims of an extended version of GEM-SURF that is currently under development.

c. High-resolution experiments

In view of the identified sensitivity of GEM-SURF’s forecast skill to resolution, the objective of the high-resolution experiments in this section is to explore the limitations of the model’s performance with respect to the resolution value. The simulations were performed over the three smaller domains shown in Fig. 1a with the grid spacing set to 250 m. It is noted that the above grid spacing lies in the middle of the range of the resolutions of the geophysical fields (Table 1). The integrations cover as before the period from 1 September 2008 to 30 November 2008. Figure 6a shows the mean diurnal evolution of the surface wind forecast from the small-domains experiment. For increase in resolution from 15 to 1.0 km there is reduction in wind speed for each increment in resolution, as in the large domain experiments. By contrast, for a further increase in resolution to 250 m, the forecast wind does not correspond to a negative speed offset relative to the 1.0-km forecast wind. Rather, it represents a rise in wind speed of different magnitude at different times of day. The 250-m simulation also gives an amplitude of the daytime rise in wind speed that is closer to the observed.

Figures 6b and 6c show the corresponding diurnal curves of RMSE and STDE of the forecast wind. A substantial reduction in value is seen as resolution increases, from a daily mean RMSE of 1.9 m s$^{-1}$ at 15 km to an RMSE of 1.5 m s$^{-1}$ at 250 m. It can be seen that the improvement in skill associated with the increment from 1.0 km to 250 m is evident during daytime and less evident during the rest of the day. It is noted that because of the small number of validating stations the spread around the RMSE curves in these experiments is larger than for the large domains (Fig. 5b).

In the bias curves (Fig. 6d), it is seen that the tendency toward a smaller daily mean bias value as resolution increases from 15 to 1.0 km reverses sign between 1.0 km and 250 m. Accordingly, Fig. 6a shows that the surface wind from the 250-m simulation is stronger than the forecast wind from the 1.0-km simulation. In the $Z_0$ change with resolution averaged for the 10 Environment Canada sites of the small domains (Table 3) a rise in the $Z_0$ value from 15 to 1.0 km is seen, as in the case of the large domain, followed by a drop in $Z_0$ from 1.9 m for the 1.0-km resolution to 0.5 m for 250-m resolution. The latter implies a smaller drag on the predicted wind of the 250-m simulation than that on the wind of the 1.0-km simulation.

The reversal of the tendency of $Z_0$ at a certain resolution can be understood by considering the relative impact at that resolution of two competing effects. The $Z_0$ value represents the parameterized influence of the subgrid-scale variability including orographic and vegetative roughness elements. As resolution increases, the smoothing of the subgrid-scale roughness becomes less, leading to a rise in $Z_0$. On the other hand, as resolution increases the number of roughness elements parameterized into $Z_0$ is reduced, which leads to a decrease in the $Z_0$ value. In view of the above, the maximum in the model’s forecast skill should be at a resolution equal to that of the geophysical datasets since then all roughness elements present in the datasets are explicitly represented and the error associated with the $Z_0$ parameterization is minimized. The GEM-SURF simulations with the 250-m grid spacing suggest that the above may be true. However, in view of the limited number of available observing stations for the evaluation of the 250-m simulation results, the need to validate this hypothesis with a larger number of observing stations is evident. The above is the subject of current investigation.

4. Conclusions

The Canadian external land surface modeling system GEM-SURF was integrated over a set of nested domains covering Canada’s eastern provinces for the period from 1 September 2008 to 30 November 2008. In the prospect of using the system as a tool for real-time surface wind prediction its skill in generating reliable forecasts of surface winds in comparison to the Canadian operational models was evaluated. Four different resolutions were examined, ranging from the resolution of the driving operational model with 15-km grid spacing down to 250-m grid spacing, a resolution similar to that of the geophysical datasets providing information on surface characteristics.

The results showed that when run in the above range of resolutions the GEM-SURF predictions 1) outperform those of the driving model and 2) progressively improve with increasing resolution. The improvement of the performance with resolution is assigned to the additional detail in the representation of the surface characteristics for each increment toward smaller grid spacing, and also to the improved downscaling that the forcing fields
FIG. 6. (a) Mean diurnal evolution of the autumn STDE of the GEM-SURF forecast (color) and the observed (black) surface wind speeds averaged over the 10 stations situated within the Anse-à-Valleau, Carleton, and Baie-des-Sables integration domains; winds from the 15-, 2.5-, and 1.0-km and 250-m runs are in blue, green, red, and magenta, respectively; the abscissa shows the local hour of day and the ordinate gives the corresponding variable in wind speed units (m s\(^{-1}\)). (b) Mean diurnal evolution of the autumn RMSE of the GEM-SURF forecast winds over the 10 observing stations; the shaded zones around each curve represent error margins for a 95% confidence interval, obtained with the bootstrap method with a trial number of 1000; color convention is as in (a). (c) Mean diurnal evolution of the autumn STDE (dashed lines) for the 15-, 2.5-, and 1.0-km and 250-m runs, averaged over the 10 observing stations. (d) As in (b), but for the wind error bias.
undergo as a result of the more accurate representation of orography. The maximum model skill is seen at the highest resolution of the examined range. The latter is comparable to the resolution of the geophysical datasets of orography, vegetation, land–water mask, and soil type, thus suggesting that the quality and the resolution of the representation of the surface characteristics are critical for the improvement in the model’s performance.

With regard to the dependence of error on the magnitude of the wind speed, a small rise in error is found with increase in wind speed, as are overestimates of the observed speeds by the forecast for very low winds between 0 and 3 m s\(^{-1}\) and underestimates for speeds higher than 3 m s\(^{-1}\). Analysis of the behavior of the wind error bias with increasing resolution showed a tendency from a larger to a smaller positive daily mean value that represents a decrease in the forecast wind speed for each incremental rise in resolution. A close correspondence was found between, on one hand, the change in the bias value with resolution and, on the other, the change in the value of the roughness length \(Z_0\). The latter shows a rise in value in the range from 15 to 1.0 km, indicative of a strengthening of the drag force exerted on the predicted wind with increasing resolution. However, statistical downscaling tests based on the diagnosed bias showed that the improvement in model performance with increasing resolution is associated not only with a reduction of the systematic error represented by the bias, but also with a reduction in the random error.

With respect to the sources of error, inaccuracies in the forcing fields, deficiencies in the quality of the geophysical datasets, and errors related to the process of their interpolation onto the GEM-SURF grid are expected to contribute to the surface wind forecast error. The outcome of model parameterizations that are strongly dependent on the geophysical data, like the \(Z_0\) parameterization and the parameterizations associated with the hydrological–soil processes, is also affected by the aforementioned error sources. Sensitivity studies with geophysical fields different from those used in this study are needed to quantify the relative impact of each factor on the surface wind forecast error. Along these lines, exploration of alternative interpolation/remapping methods onto the model grid and on the observing sites would be of significant value, especially in cases when the density of the observing network, the resolution of the geophysical datasets, and the resolution of the GEM-SURF domain are significantly different.

In addition to the efforts for the improvement of aspects related to the geophysical data fields that pertain to all surface variables, emphasis for the surface wind predictions in particular is placed on the need for investigation for an improved formulation for the stability functions of heat and momentum \(\Phi_h\) and \(\Phi_m\). Given the impact that the latter have on both the wind and the surface temperature forecasts and the higher quality of the temperature forecasts relative to the winds seen in these experiments, it is considered that the simulated, on the basis of \(\Phi_h\) and \(\Phi_m\), effect of the near-surface turbulent fluxes, should be reevaluated so that the surface wind scores are improved and the satisfactory current scores of the temperature forecasts retained. The weight put on maintaining the surface temperatures at realistic levels introduces an enhancement/weakening of the near-surface turbulent fluxes in the cold/warm part of the day in the present formulation and causes the overestimates/underestimates in the surface winds seen in these experiments at the respective times.

The region of eastern Canada over which the experiment was performed covers a wide range of physiographic conditions in terms of elevation, land–water distribution, vegetation, land use, and soil-type characteristics. In view of this diversity it is suggested that the presented GEM-SURF results are applicable to a large number of geographical regions in the midlatitude zones. The system’s performance over regions with more extreme characteristics, like alpine terrain, extensive croplands, or semiarid climate zones, remains to be examined. Similarly, the validity of the results in lower and higher temperature conditions needs to be assessed for application of the results to seasons other than autumn. Last, in view of the encouraging results from the high-resolution experiment, the minimal cost in computer resources involved in operating the system at high resolutions, and the continuous flow of geophysical information in the form of satellite products of high quality and resolution, it is suggested that exploration of the model’s skill through experimentation with highest-resolution geophysical data is a promising direction to pursue in future research efforts.

Acknowledgments. This research was funded by the Canadian Government ecoETI program.

REFERENCES


——, R. Brown, J. Mailhot, B. Bilodeau, and L.-P. Crevier, 2003b: Operational implementation of the ISBA land surface scheme


