Very High Resolution Coupled Weather and Wind Power Modeling

L. Treinish, J. Cipriani and A. Praino
IBM Thomas J. Watson Research Center
Yorktown Heights, NY, USA
{lloydt,jpcipria,apraino}@us.ibm.com

J. R. Dorronsoro and J. Díaz
Instituto de Ingeniería del Conocimiento and
Universidad Autónoma de Madrid
Madrid, Spain
{julia.diaz,jose.r.dorronsoro}@iic.uam.es

Other Presentations of Related Work

Conference on Weather, Climate, and the New Energy Economy:
- 6.3 Enabling Advanced Weather Modelling and Data Assimilation for Utility Distribution Operations
- 8.1 Outage Prediction and Response Optimization (OPRO)
- 10.1 Improvements in short-term solar energy forecasting – Thursday morning
- 10.2 Two methods in improving onshore wind forecast – Thursday morning

Conference on Numerical Weather Prediction:
- 9.3 Ensemble Kalman Filter (EnKF) Assimilating the Dropsonde Observations to Reduce the Forecast Track Error of Typhoon Soulik (2013) Based On the Cloud-resolving Model
- 13.2 Recent Advances in High-Resolution Operational NWP, Utilizing WRF-ARW – Thursday morning

Conference on Artificial and Computational Intelligence and its Applications to the Environmental Sciences:
- J3.2 A multi-scale solar energy forecast platform based on machine-learned adaptive combination of expert systems – Wednesday morning

Conference on Climate Variability and Change:
- 8C.4 Simulation of the temporal and spatial characteristics of diurnal rainfall cycle over Borneo

Symposium on Advances in Modeling and Analysis Using Python:
- 3.5 A Python-Based Automatic Data Aggregation Framework for Hydrology Models

Superstorm Sandy and the Built Environment: New Perspectives, Opportunities, and Tools:
- 873 Forecast Performance of an Operational Mesoscale Modeling System for Post- Tropical Storm Sandy in the New York City Metropolitan Region

Conference on Probability and Statistics in the Atmospheric Sciences
- 4.2 Customized Verification Applied to High-Resolution WRF-ARW Forecasts for Rio de Janeiro
- 6.5 Statistical forecasting of rainfall from radar reflectivity in Singapore

Symposium on the Urban Environment
- J12.2 High-Resolution, Coupled Hydro-Meteorological Modelling for Operational Forecasting of Severe Flooding Events in Rio de Janeiro

Precision Wind Power Forecasting via Coupling of Turbulent-Scale Atmospheric Modeling with Machine Learning Methods

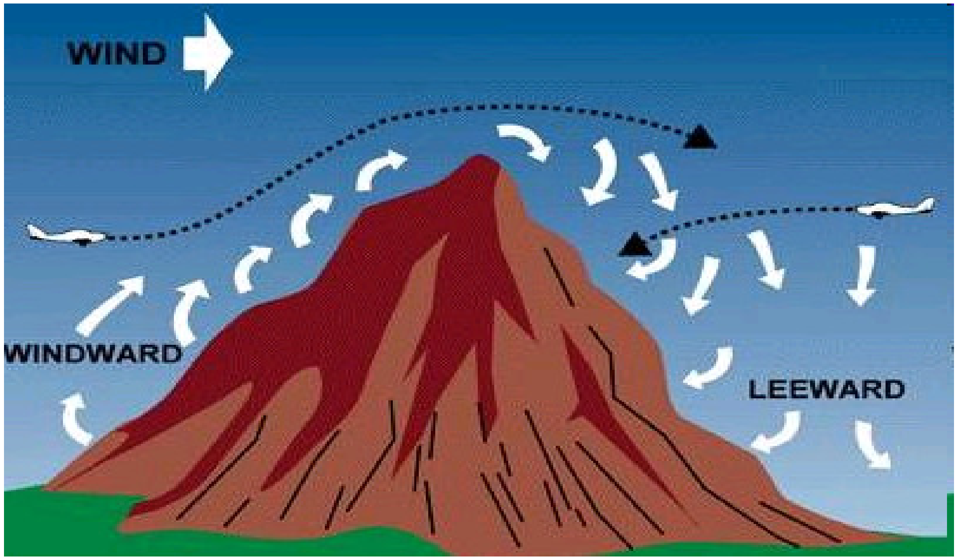
- **Motivation and background**
- **Approach**
- **Preliminary results**
- **Project plans and status**

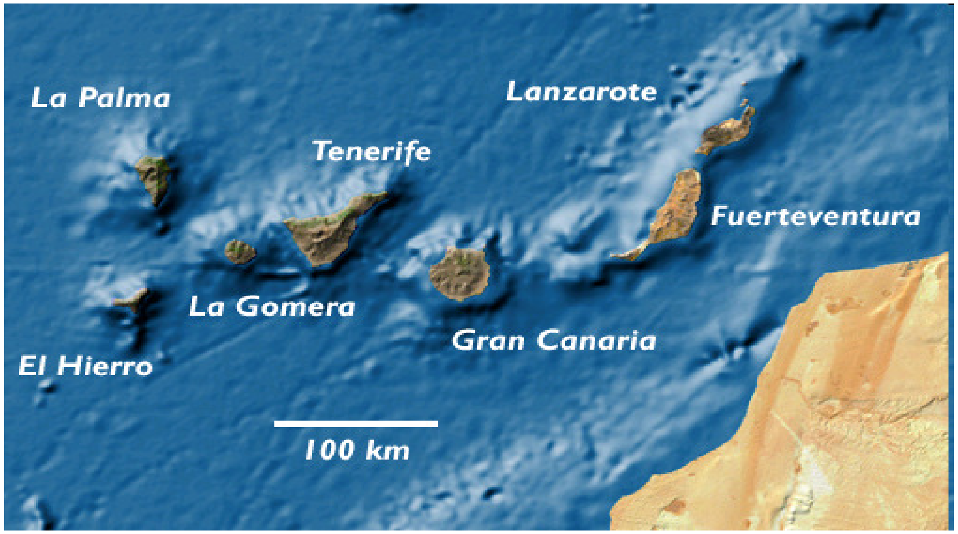
Motivation and Background

- **Wind power intermittency creates significant barriers to expanding utilization**
  - Ramp events
  - Spinning reserve
- **Better forecasting and optimized economic dispatch can alleviate these barriers**
  - Ensemble forecasts
  - Stochastic programming
  - Dynamic reserves
- **Challenges are greater for isolated systems such as on islands**
  - No grid interconnection with larger systems

Canary Islands

- Isolated system across the archipelago (7,493 square km of land area)
- Red Eléctrica de España: 45 wind farms, located on 5 of the 7 islands with ~142 MW aggregated capacity with a wide variety of equipment (327 turbines)
- Complex topography leads to turbulent flow, especially along the coastlines
  - –Peaks up to 3500m, inducing vortices
- Large power output variability as a result of ramp events
  - –For example, 7 November 2010: 61% variability in a four-hour time span on Gran Canaria
  - –Impacting reliability, electricity generation
- Ramp events are poorly predicted
  - –NWP-based forecasts do not capture flow
  - –Machine learning and statistical methods are brute force and lack good training sets



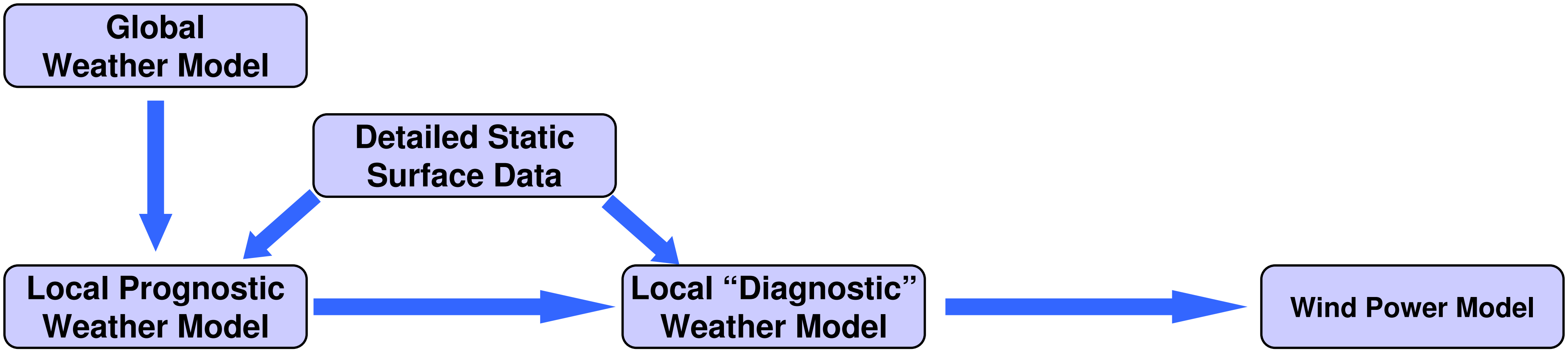


Approach

- Given the geography of the archipelago, and the spatial distribution of the individual turbines, turbulence-scale modeling becomes essential to capture the flow
- Introduce large eddy simulations (LES)
  - Capture planetary boundary layer (PBL) effects
  - Momentum, heat, moisture flux terms become critical
  - High-temporal resolution required to capture transients (output every five minutes)
- Retrospective analysis of critical ramp events
  - Many numerical experiments to enable effective model configuration
- Balance detail vs. performance
  - Horizontal/vertical resolution vs. time step, subject to CFL stability criteria

    $|U|*dt/dx <= 1$

Approach to Coupled Weather and Power Modelling



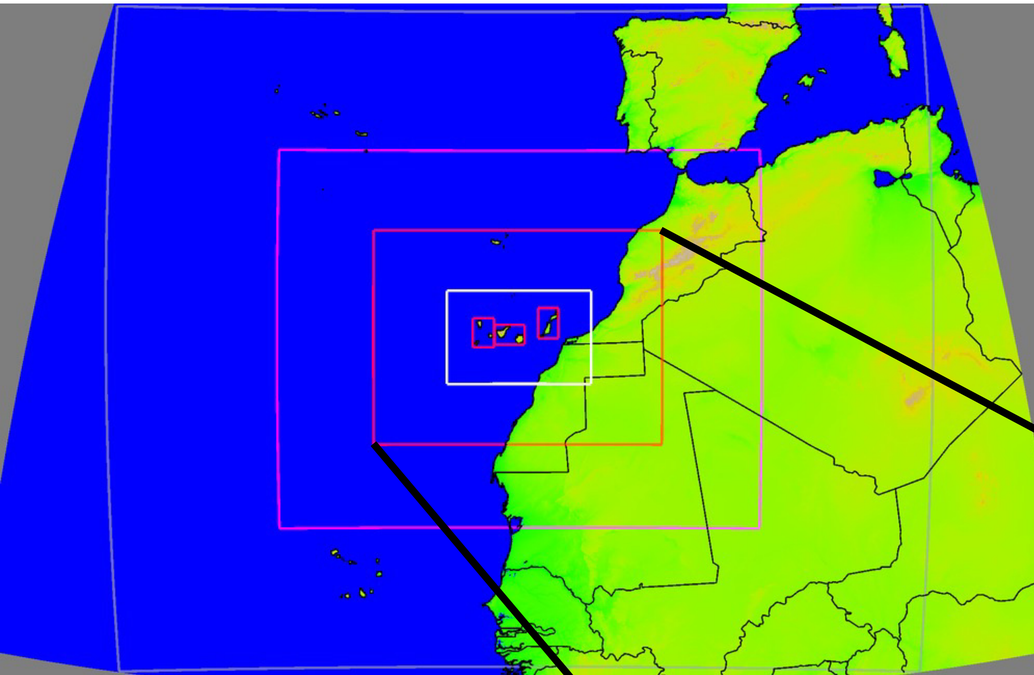
Local Prognostic Weather Model
- 1-2 km horizontal resolution
- 20-1000m vertical resolution
- Full dynamics and 3d physics

Local "Diagnostic" Weather Model
- 400-800m horizontal resolution
- 20-1000m vertical resolution
- Turbulent-scale dynamics and more limited physics

Weather Model Configuration: WRF-ARW 3.3.1

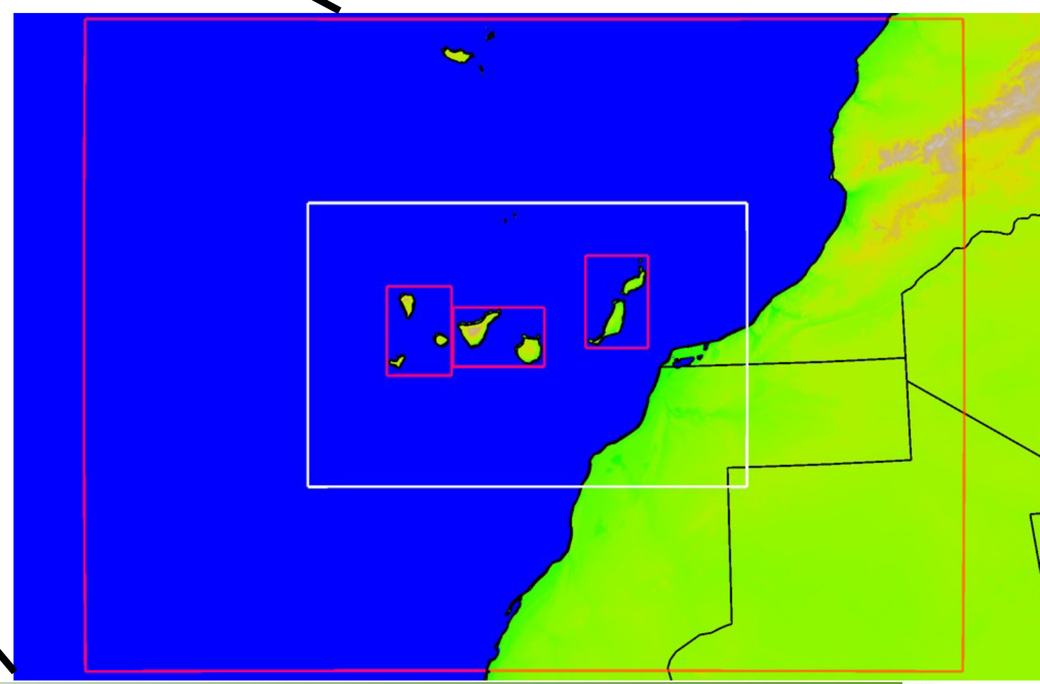
- 50 vertical levels with ~10-15 in the planetary boundary layer to ensure capturing of orographic effects
- 24 hour runs initialized at 0 UTC
- NOAA GFS for background and lateral boundary conditions
- SRTM-based model orography (90m)
- MODIS-based land use data
- 1km-resolution JPL SSTs
- Four 2-way nests at 54-km (87x70), 18-km (151x118), 6-km (268x199), 2-km (358x244) focused on the Canary Islands
  - WSM 5-class single moment microphysics, RRTM long wave radiation, GSFC short wave radiation, YSU PBL, NOAH LSM, Kain-Fritsch cumulus physics
- Three one-way LES domains embedded within domain four at 666.67m resolution (178x244, 250x163, 172x253)
  - WSM 5-class single moment microphysics, new GSFC long and short wave radiation, LES PBL, NOAH LSM, explicit cumulus physics
- Data assimilation is not feasible given the lack of a comprehensive observing system

Weather Model Configuration

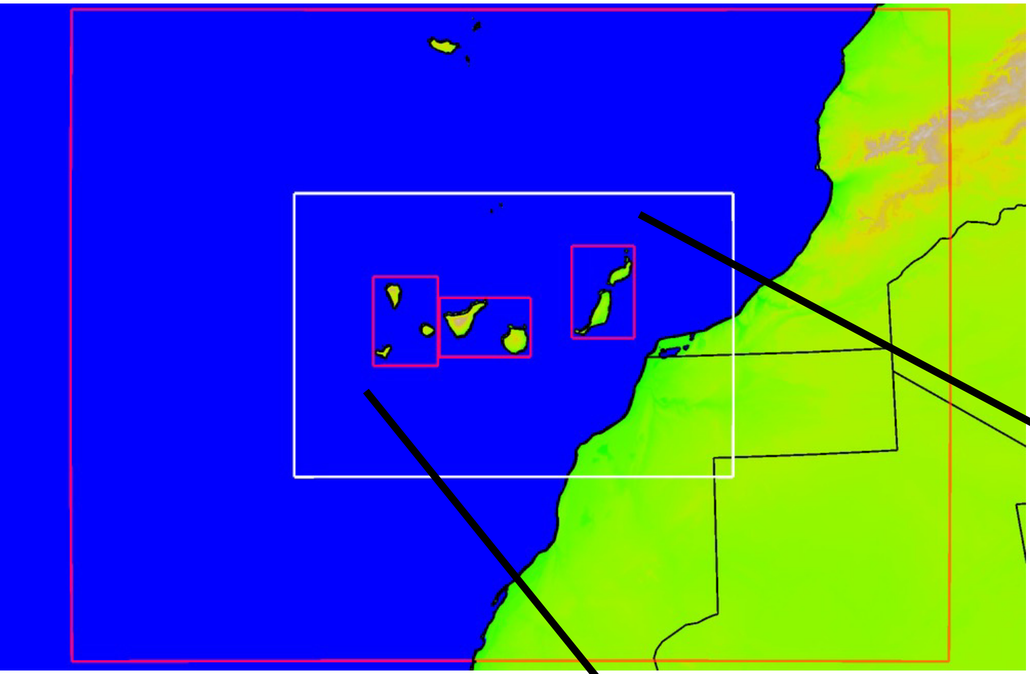


Four 2-way telescoping nests at 54, 18, 6 and 2 km horizontal resolution driving three, independent one-way LES nests at 667m resolution focused on the Canary Islands

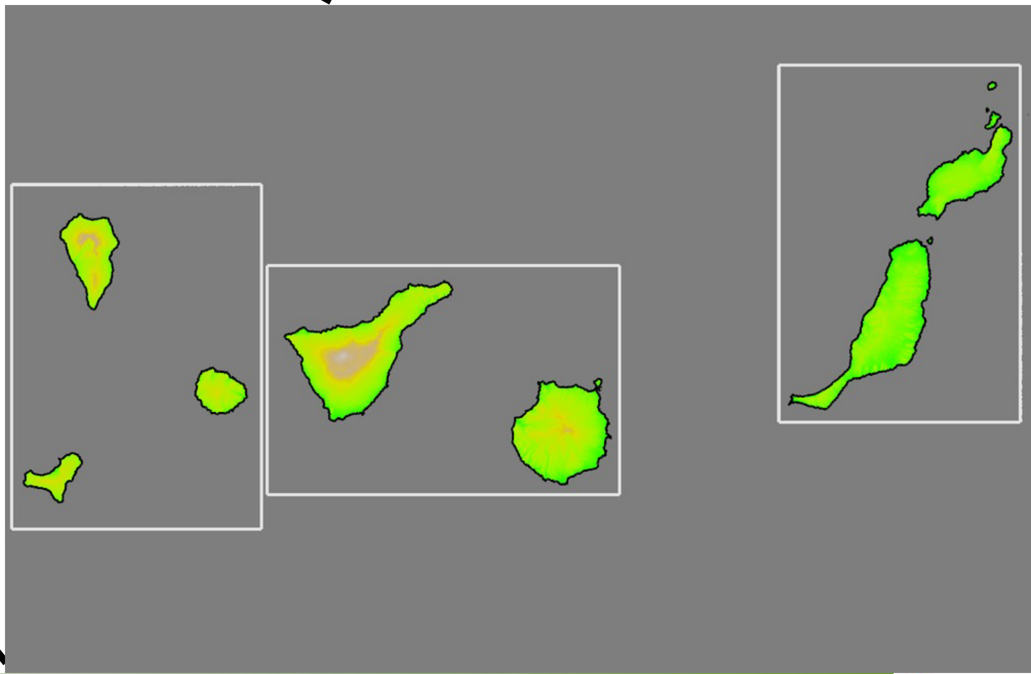
50 vertical levels with 10 to 15 in the planetary boundary layer



Weather Model Configuration



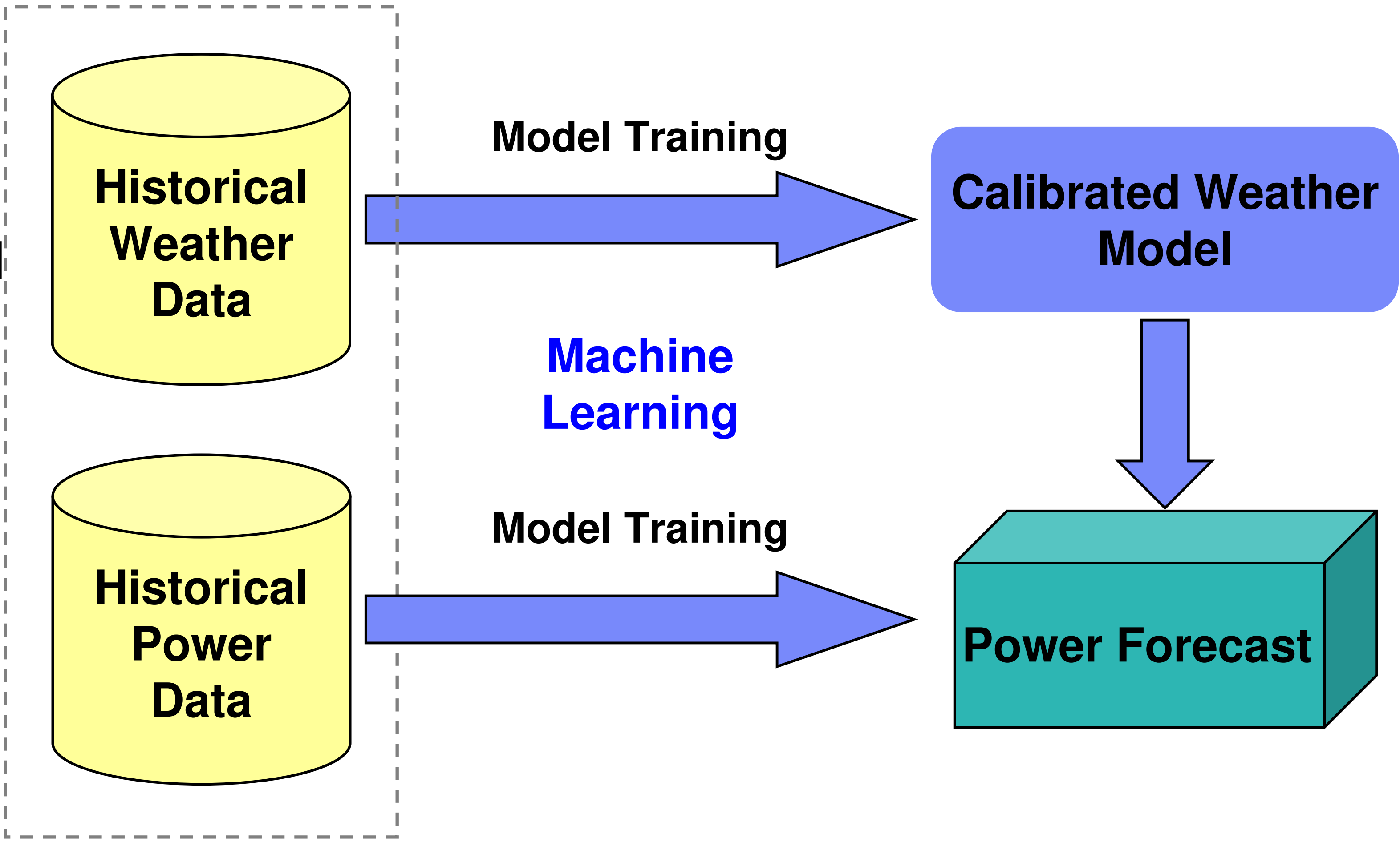
Four 2-way telescoping nests at 54, 18, 6 and 2 km horizontal resolution driving three, independent one-way LES nests at 667m resolution focused on the Canary Islands



50 vertical levels with 10 to 15 in the planetary boundary layer

Coupled Weather and Power Modelling

- Build predictive model from historical weather forecasts and power, and related data
  - –Wind farm power
  - –Turbulent flow
  - –Wind farm locations and characteristics



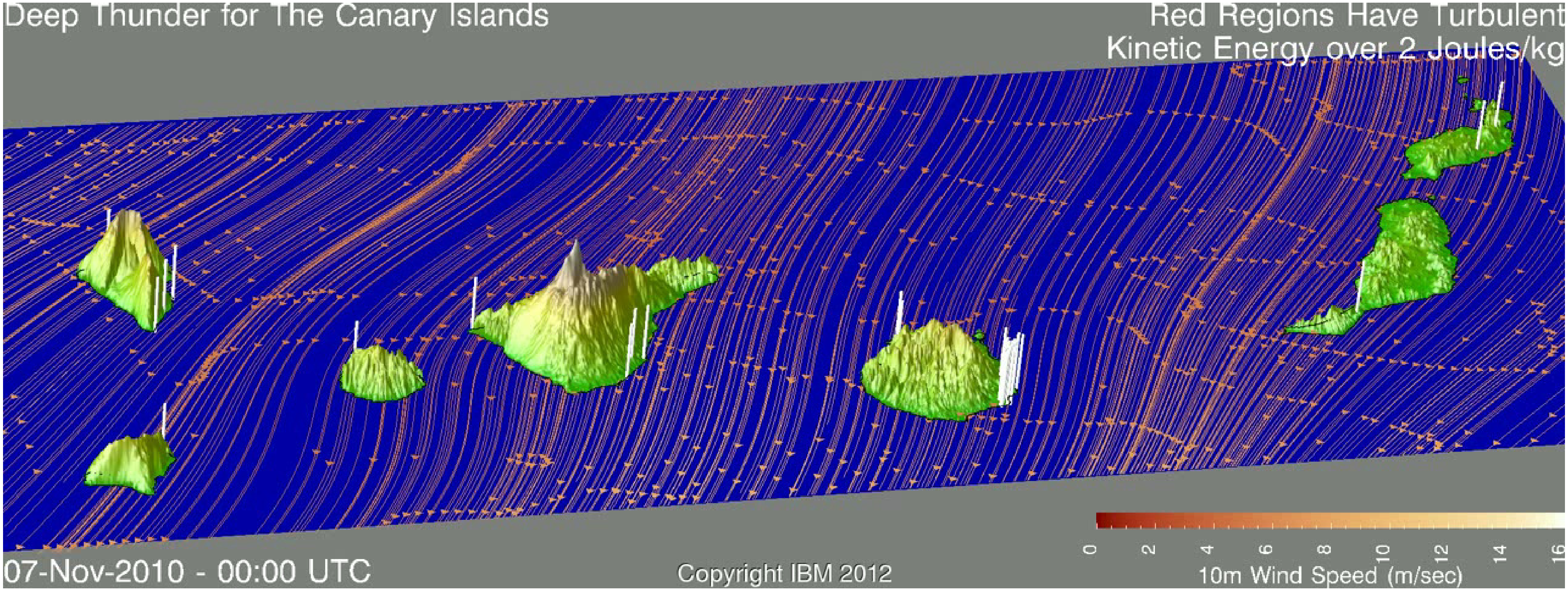
Approach

- Create a targeted hindcast database as a training set for the machine learning algorithms, based on the atmospheric physics
  - Starting with 19 ramp events throughout 2010 and 2011
  - Goal: enable a six-month continuous period (daily)
- Avoid brute-force machine learning approach
  - Use only data that relate to energy extraction process
  - Volumetric Turbulent Kinetic Energy (TKE), absolute vorticity, 3-vector wind fields
  - Derived surface gusts and Clear-air Turbulence (CAT) index (two-dimensional)

Additional Challenges: Verification

- **Insufficient weather observations: 7 stations across the archipelago with only hourly reporting**
  - May miss the transient events
  - 5-minute interval data from hindcasts unverifiable
- **Power data are "limited"**
  - Hourly percentage increase/decrease, aggregated over each island
  - No power curve information for each turbine or farm
  - Only two years available: 2010-2011

Example Results: 7 November 2010 Ramp Event 0300 - 0700 UTC



Animation of 10m winds and TKE (red isosurfaces at 2J/kg) indicative of turbulent flow

- **The 10m wind streamlines above the water are derived from the 2km nest (4)**
- **The isosurfaces are derived from the three 666.67m nests (5)**
- **The terrain of each island is shown**
- **The location of each of the 45 wind farms are marked with a white pole whose height corresponds to the blade extent for the deployed turbines**

Machine Learning (ML) for Power Forecasts

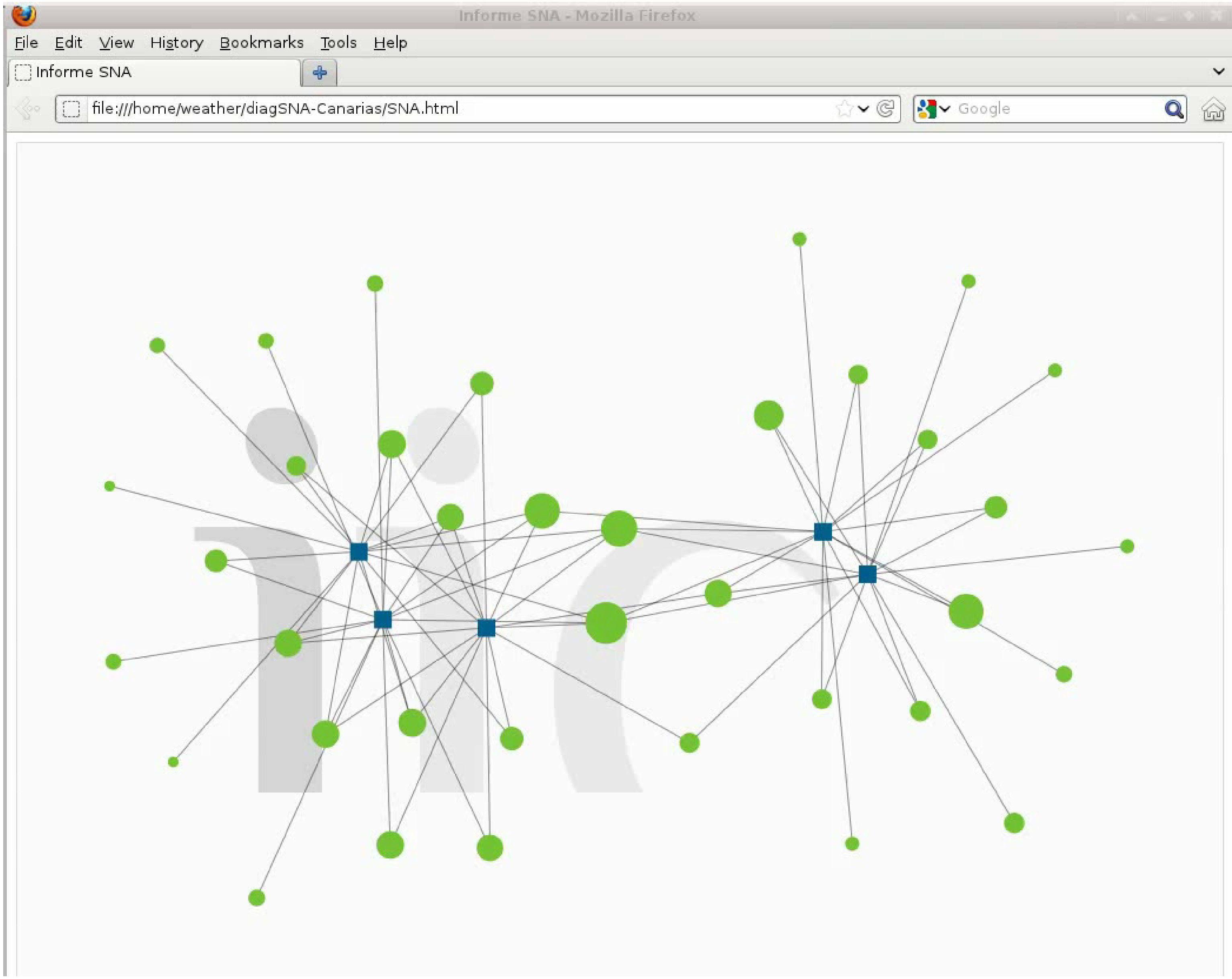
- Transform NWP output into energy forecasts
- Currently ML-based energy forecasts are derived from numerical patterns obtained from NWP outputs derived at the synoptic scale (from ECMWF)
- This approach implies both a large increase of the NWP pattern dimension and of sample size, which goes from eight to at least 24 patterns per day
- Sample sizes and dimension will have the same order of magnitude, contrary to the ML rule of thumb of sample size being an order of magnitude greater

Machine Learning (ML) for Power Forecasts

- **The very fine resolution data from the turbulent-scale NWP input can be extremely large for ML models**
  - This clearly precludes the application of non-linear models, as their time complexity would be prohibitive
- **It has been observed in other fields that simple linear models can yield good results for problems with large dimensional inputs**
  1. Linear Support Vector Regression (SVR), which uses the so-called hinge-loss that penalizes only forecast errors above a certain tolerance
  2. Elastic Net and Lasso methods
  - Both combine a square error function with an L1 regularization penalty term
  - Elastic Net also adds a quadratic penalty, as done in ridge regression
- **Both approaches present two important properties**
  1. The models are built solving a convex optimization problem and, thus, have a unique minimum value
  2. The hinge loss of SVR and the L1 regularization of Lasso and Elastic Net result in sparse final models with many zero coefficients that enable a fast application to new data and also can be exploited for ranking the predictive NWP variables

Social Network Analysis (SNA) Graph of Results

- Each island is a blue rectangle
- Each meteorological variable is a green circle with the corresponding weights
- A specific island can be highlighted by clicking on a rectangle
- The weight of a circle corresponds to the aggregation of weights for that variable in all the separate models per island
- Each island shows the weight of the variable for that model
- Gran Canaria and Tenerife are in the same cluster: the most predictive variables are the same
- A second cluster is for the others islands



Machine Learning (ML) for Power Forecasts

- The forecasts produced by the combination of the turbulent-scale NWP forecasts and the linear models will be compared with those obtained using ECMWF forecasts at a 0.25 degree resolution

- The linear models using the turbulent-scale NWP have been built using a comparatively short two-month training period

- The ECMWF forecasts were used with the non-linear models built using a 12-month training period

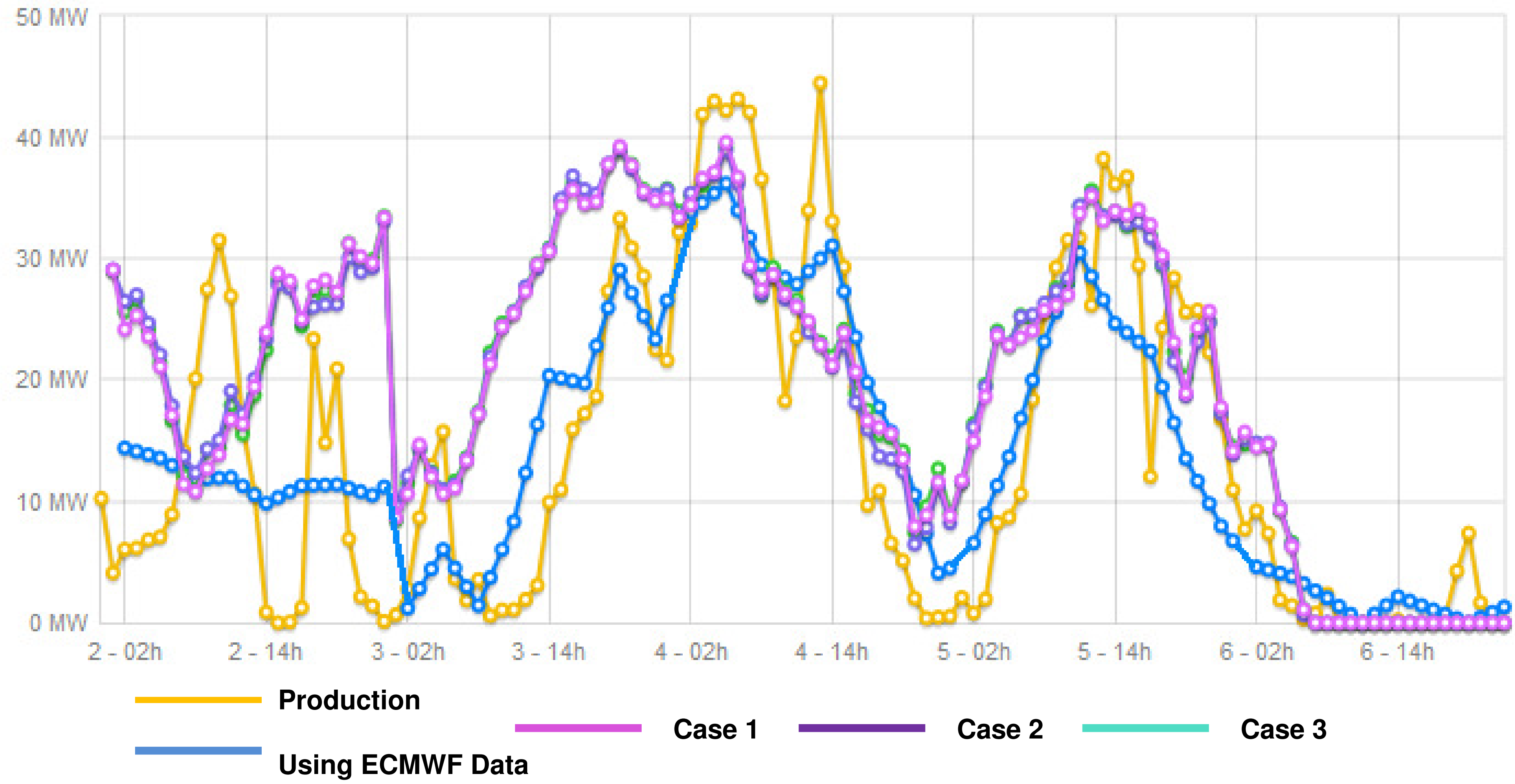
- 10 different instances with different parameterizations are shown with hourly output for three different sample sizes

Predictions for Gran Canaria

Prediction	ECMWF	Case 1	Case 2	Case 3
Resolution	0.25	Mean: 0.006	Mean: 0.006	Mean: 0.006
Grid points	207	126441	126441	126441
Variables	4	71 (10 levels)	71 (10 levels)	71 (10 levels)
Step	3-Hourly	30 minutes	5 minutes	15 minutes
Historical data	18 months	2 months	2 months	2 months
Space required	Total: 500 MB	Total: 600 GB	Total: 600 GB	Total: 600 GB
Execution time	20 minutes	1.3 hours	> 10 hours	2.2 hours

Predictions for Gran Canaria

Test Period: 10/02/2010 - 10/06/2010

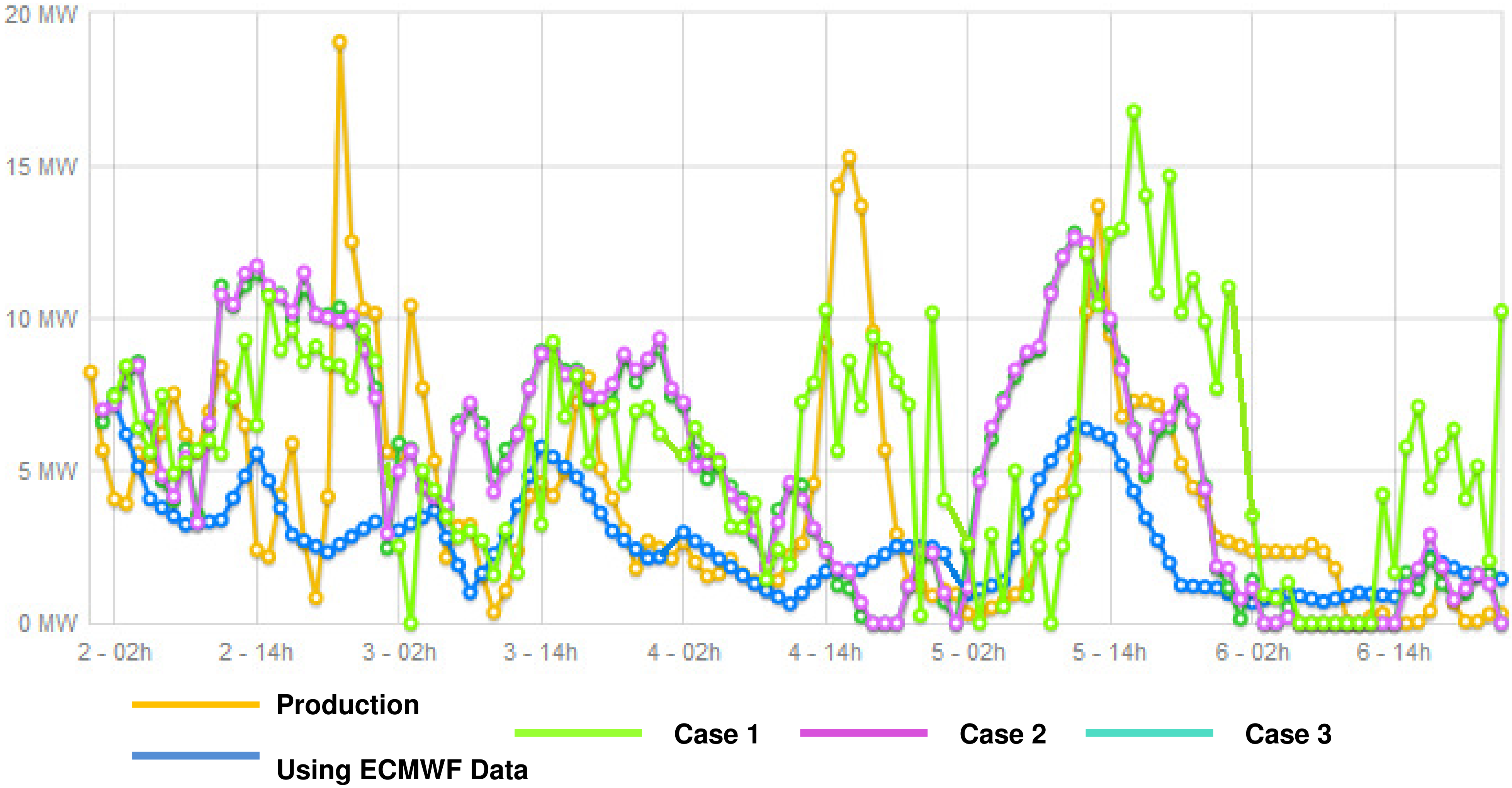


Predictions for Tenerife

Model	ECMWF	Case 1	Case 2	Case 3
Resolution	0.25	0.1	Mean: 0.006	Mean: 0.006
Grid points	207	506	126.441	126.441
Variables	4	38	71 (10 levels)	71 (10 levels)
Steps	3-Hourly	Hourly	5 minutes	30 minutes
Historical data	18 months	2 months	2 months	2 months
Space required	Total: 500 MB	Total: 600 GB	Total: 600 GB	Total: 600 GB
Execution time	20 minutes	25 min	4 hours	54 min

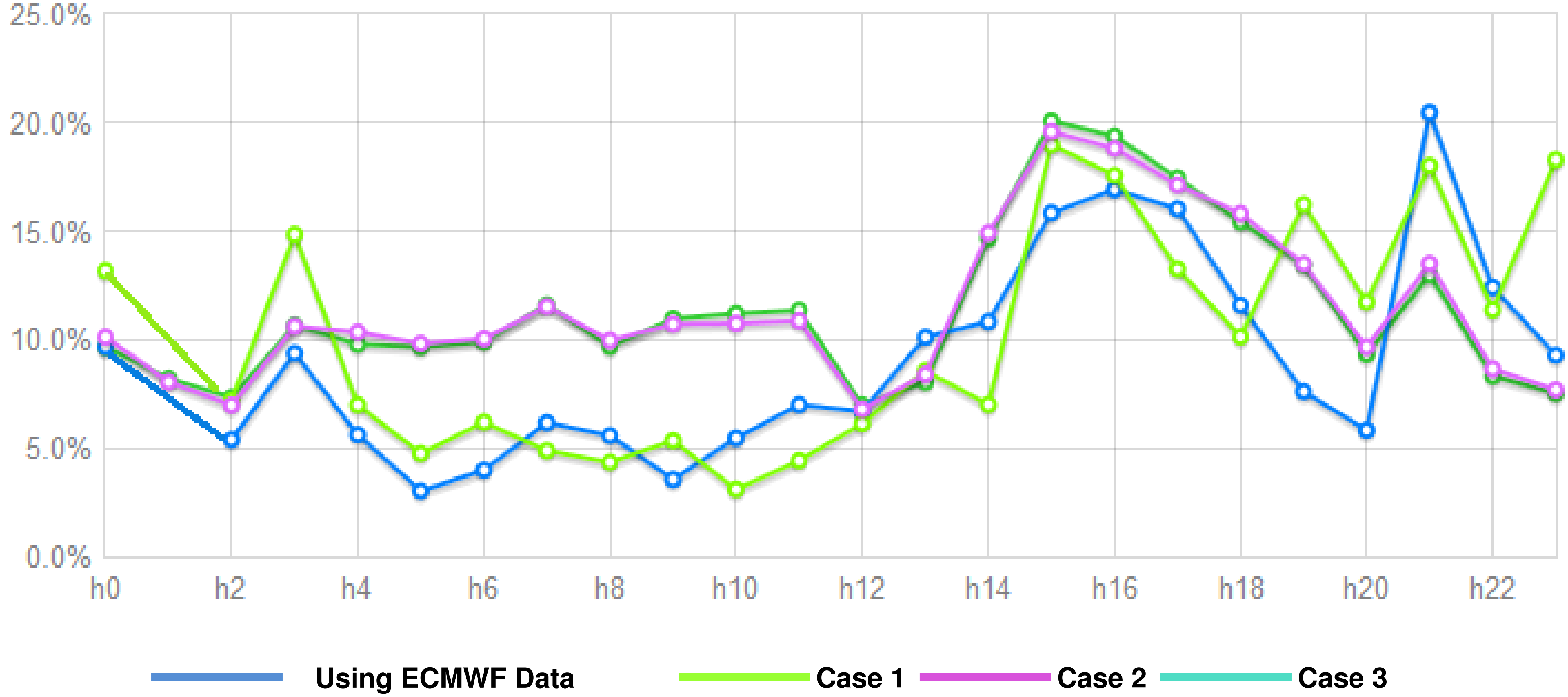
Predictions for Tenerife

Test Period: 10/02/2010 - 10/06/2010



Comparison of Prediction for Tenerife to Production (Root Mean Square Error)

Test Period: 10/02/2010 - 10/06/2010



Conclusions

- Model development and optimization were more complex than originally estimated
- Current method uses 18 months of ECMWF data
- Methods using turbulent-scale NWP have been trained with only 2 months of data
- Despite the difference in the training periods, the quality is similar
- In some periods, the error is lower with the turbulent-scale NWP for some islands
- Significant potential for the application of fine-resolution NWP models for forecasting wind energy over islands or, more generally, isolated renewable energy systems

Future Work

- Expand size of training set, retrain the models and evaluate the results, including additional metrics
- Evaluate other potential NWP configurations
- Determine an optimal balance between the current temporal (5 minute) and spatial (667m) resolution to reduce the computational cost for the ML methods
- (Limited) verification of weather hindcasts

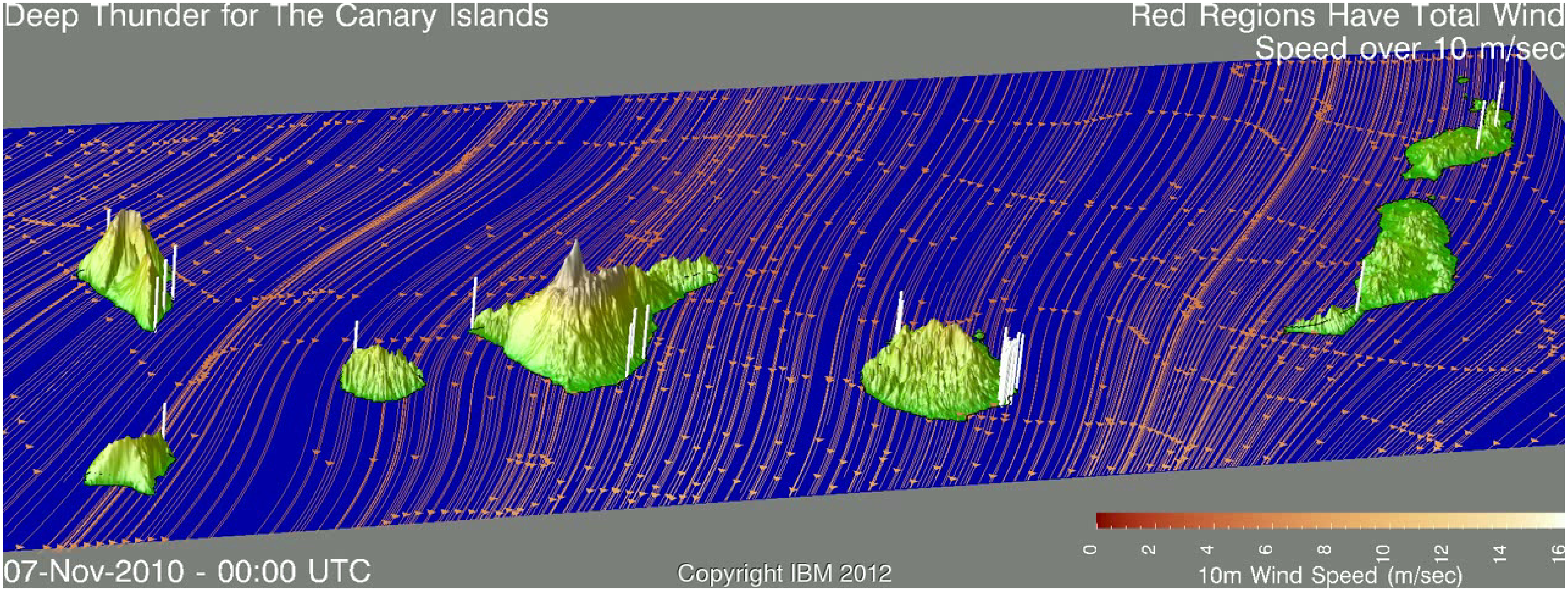
Backup

Slides

Computational Issues

- **Ordinarily, significant resources required to enabling LES forecasts in a production environment over a large domain**
  - Some effort to address the practicality using a modest HPC system
  - Effort to build training set and potential for operational use
- **Optimized for a cluster of ten 32-way Power7 nodes, each with 256GB memory with a DDR Infiniband interconnect**
  - Four 2-way nests (to 2km) run 24 hours in 50 minutes on six nodes
  - Three 667m LES nodes run in parallel, each using three nodes requires about 100 minutes as an NWP post-process (1-way nests)
  - End-to-end processing is about 3.5 hours per 24-hour simulation
  - Six-month climatology requires ~630 hours of compute time
- **Each run generates 180 GB of data (uncompressed)**
  - Most of the data are not relevant to drive machine learning
  - Four 2-way nests are cheaper to recompute than store
  - Only store fields related to turbulent flow in lower part of the boundary layer, resulting in a six-month climatology ~1 TB in size

Preliminary Results: 7 November 2010 Ramp Event 0300 - 0700 UTC



Animation of 10m winds and [u,v,w] (red isosurfaces at 10m/sec) indicative of turbulent flow

- **The 10m wind streamlines above the water are derived from the 2km nest (4)**
- **The isosurfaces are derived from the three 666.67m nests (5)**
- **The terrain of each island is shown**
- **The location of each of the 45 wind farms are marked with a white pole whose height corresponds to the blade extent for the deployed turbines**