A methodology to obtain the best WRF configuration for operational forecasts. Application over a coastal region in southern Spain

RAÚL ARASA, IGNASI PORRAS, ANNA DOMINGO-DALMAU, MIQUEL PICANYOL*, M^{A} ÁNGELES GONZÁLEZ, JÉSICA PIÑÓN

Meteosim S.L., Barcelona, Spain.

BERNAT CODINA

Department of Astronomy and Meteorology, University of Barcelona, Barcelona, Spain.

1. Introduction

Meteorological/NWP models have a wide range of options to set up: physics options, dynamics options, horizontal model resolution, number of vertical layers and density, domains architecture and nest down options, data assimilation, time-step, spin-up time, etc. It is a fundamental factor when configuring a model the selection of the parameterizations and options that are used (Stensrud 2007). The best combination for one region (Krieger et al. 2009) is not necessarily applicable to another. Scientific community has drawn up guides and recommendations on the use of meteorological models (Warner 2011; Denby et al. 2008; Wang 2014; Dudhia and Wang 2014) that modelers may consider previously to simulate. In this sense, authors have defined and established an own standard methodology that could be applied in any region and allow to improve meteorological forecast for operational purposes.

We focus our attention on the Port of Huelva (Huelva), a region in southwestern Spain. This work aims to investigate the best configuration of a meteorological model that allows to reduce the uncertainty and, therefore, to increase the confidence level of the forecasts. We have used the WRF model to obtain the meteorological forecasts and we have defined a procedure to calibrate the model in a customized way for Huelva but it can be applied to any region. These forecasts will be used as an early warning system and will allow improving the risk management associated to daily activities, and particularly, to manage more efficiently the atmospheric pollution generated as consequence of aggregate handling and storage piles associated to the treatment of different solid materials, providing a better air quality for the region.

Corresponding author address: Miquel Picanyol, Meteosim S.L., Barcelona Science Park, Baldiri Reixac 10-12, 08028 Barcelona, Spain.
E-mail: mpicanyol@meteosim.com

2. Methodology

a. Studied area, simulation domains and episodes selected

The methodology defined to obtain the best WRF configuration has been applied over Huelva, in southwestern Spain. Air quality levels achieved and the risk management in a complex harbor located very near of the population made remarkable the implementation of a very accuracy meteorological model in the zone.

Huelva had a population of 149,410 in 2010 and it is located along the Gulf of Cadiz coast in the mouth of the Odiel and Tinto Rivers. Huelva has a subtropical-Mediterranean climate characterized by dry and hot summers and wet and mild winters. Temperatures higher than 40 °C are usually reproduced on summers. Highest wind velocities occur during the sunset and episodes of severe gale (force 9 in the Beaufort scale) affects the region occasionally during the year.

In the city of Huelva and its metropolitan area coexist the core of the population of the province of Huelva, greenhouse zones, nature reserves (very near of Doñana Park) and one of the most important industrial poles in the south of Spain. The activity of the industrial sector is mainly characterized by the Port of Huelva, divided in two sectors: the inner port (near the city of Huelva) and the outer port, being this last the most important. Activity in the Port is associated with a high flow of loading and unloading operations and material handling. In this sense, the meteorology influences the atmospheric pollution generated by these processes and the activity, in itself, is conditioned by the meteorological conditions.

In Figure 1, we show the modeling domains used for operational forecasts over the coastal region of Huelva. The WRF model is built over a mother domain (d01) with 9 km spatial resolution, centered at 37.14°N 7.38°W. It comprises the southwestern Iberian Peninsula and north of Morocco with an important zone of sea space, and with a domain size of 927 x 900 km^2. This domain is intended to

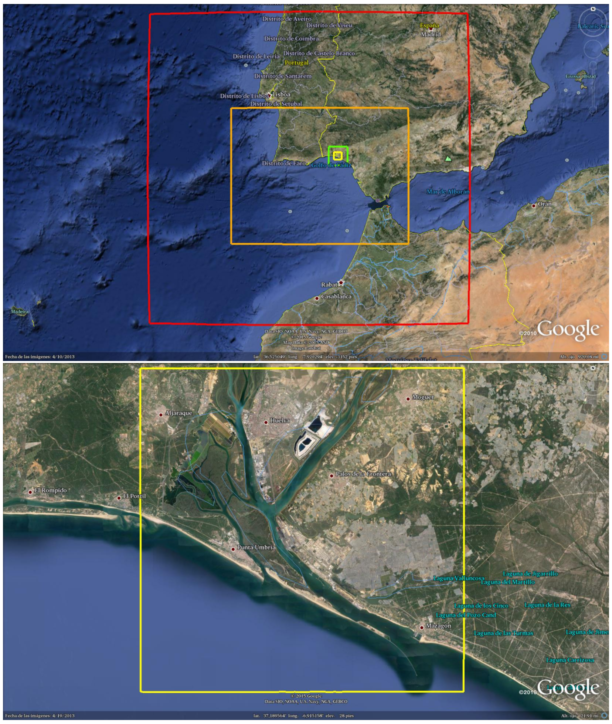

FIG. 1. Modeling domains used for the simulations: d01 (9 km), d02 (3 km), d03 (1 km) (top), and d04 (0.333 m) (bottom).

capture synoptic features and general circulation patterns. The first nested domain (d02), with a spatial resolution of 3 km, covers the south of Portugal and western Andalusia with a domain size of 507 x 390 km^2. The third domain (d03) with 1 km of spatial resolution covers Huelva, and the fourth domain (d04) with 0.333 km covers the port area of Huelva with an extension of 500 km^2 approximately.

Simulations were conducted in different periods of the year 2012 and 2013. Numerical simulations are executed for 30 hours corresponding on every day (hereinafter referred to as daily simulations) included in the period compressed between 01/01/2012 and 12/31/2013, taking the first 6 hours as spin-up time to minimize the effects of initial conditions. Different periods have been selected depending on their final use:

- To calibrate the model we have considered the months of February, May, August and October for the year 2013. These months represents the climate variability of the region, being the coldest, driest, warmest and wettest month respectively. A total of 3120 WRF daily simulations have been conducted, corresponding to 120 days and 26 different experiments.

- To validate the model we have considered two full years, 2012 and 2013, and therefore, 730 WRF daily simulations have been conducted.

- And to analyze the experiments of data assimilation we have considered an operational forecasting period, corresponding to the period compressed between 08/24/2015 and 10/24/2015. For data assimilation analysis a total of 300 WRF simulations have been conducted, corresponding to 60 days and 5 experiments.

b. Modeling approach

WRF has different parameterizations for microphysics, radiation (long and short wave), cumulus, surface layer, planetary boundary layer (PBL) and land surface as physics options. To obtain the WRF highest accuracy, it is essential to carry out a sensitive analysis of these different options by numerical experiments. In the same way, the definition of the simulation domains, spin up, vertical resolution or nesting architecture determine the accuracy, and therefore uncertainty, of WRF results (R. Arasa 2012).

The initial and boundary conditions for the operational configuration over domain d01 were supplied by the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) with an horizontal resolution of 0.25° and updated every 6 hours. To calibrate and validate the model a new parent domain has been defined (d00), covering all Iberian Peninsula, south of France and northwestern Africa. The parent domain d00 has been defined to adjust the coupling between global and mesoscale model. In this case, for the sensitive analysis initial and boundary conditions have been supplied by the NCEP/NCAR Climate Forecast System Reanalysis v2 (CFSv2) with 0.5° of spatial resolution and 6 h of temporal sampling. CFS is a better atmospheric representation than GFS because incorporates a higher amount of observations and measurements. In both cases, two-way nesting was used for the external domains (d00, d01, d02 and d03) and one-way nesting for the innermost domain (d04).

In d04, the Large-Eddy-Simulation (LES) technique has been applied. This option replaces the traditional patterns of atmospheric boundary layer using the mathematical model of turbulence originally proposed by Smagorinsky (1963) and evolved into the latest versions of WRF. This technique allows to evaluate more effectively atmospheric turbulence and allows explicitly calculate all the physical and dynamical processes that characterize the microscale, to deal well cloud formation, turbulence, heat transfer, the exchange of heat or gases fluxes, etc. This technique is considered relevant when the horizontal resolution is below 500 m.

c. Sensitive analysis and calibration

1) PHYSICS OPTIONS

First of all, we have realized experiments modifying the physics options. A total of 18 experiments have been

TABLE 3. Vertical levels experiments.

Experiment	Vertical levels
INI	30 (default configuration)
VER1	36 (15 below 1.500 m and first level at 16 m)
VER2	42 (21 below 1.500 m and first level at 8 m)

evaluated progressively, as Table 1 shows. Two of them by varying microphysics schemes, four of them by varying radiation scheme options, three by varying cumulus schemes and eight of them by varying PBL and surface layer schemes (at the same time due to model restrictions). The first numerical experiment corresponds to the default WRF options (defined as INI experiment). Secondly, microphysics experiments (MPH) are analyzed. The microphysic scheme allows to predict water phase transitions in the atmosphere and to consider snow and hail. Thirdly, longwave (LWR) and shortwave (SWR) experiments are analyzed. These schemes define radiation parameters depending on cloud cover, location, gases and aerosols in the atmosphere, time of the year, etc. Fourthly, cumulus experiments (CUM) are analyzed. Cumulus parameterization is used to predict the collective effects of convective clouds at smaller scales as a function of larger-scale processes and conditions. And finally, experiments of PBL and surface layer have been carried out (PBL). The PBL and surface layer schemes define boundary layer fluxes (heat, moisture, momentum) and the vertical diffusion process.

2) DYNAMICS OPTIONS

We have also investigated some dynamics options (shown in Table 2). We have focused on damping and diffusion. These options could improve model-top reflection of mountain waves, remove poorly resolved structures and reduce noise at model scales similar to grid-spacing.

3) NUMBER OF VERTICAL LEVELS

There are numerous papers which demonstrate that increasing the number of vertical levels is related to an improvement in the accuracy of the forecasts (Garreud and Rutllant 2003; Seaman et al. 2009). For this reason, we have started with the WRF default configuration and have increased this number to define the experiments shown in Table 3.

In Figure 2 we show a comparison between the levels for the three experiments and their distribution in sigma coordinates.

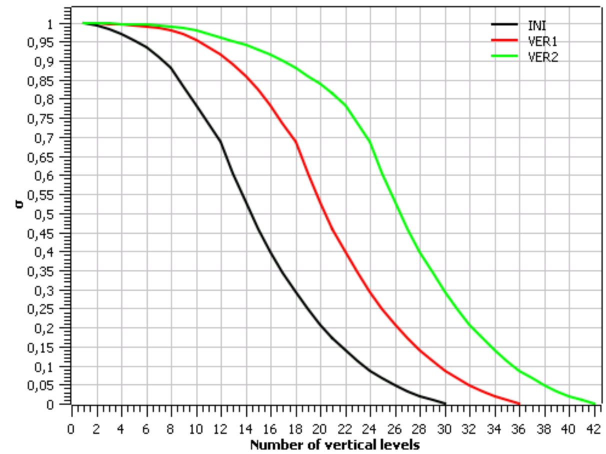

FIG. 2. Sigma level distribution for each vertical level experiment.

TABLE 4. Physiographic model databases experiments. (SRTM: Shuttle Radar and Topography Mission, ASTER: Advanced Spaceborne Thermal Emission and Reflection Radiometer, CLC2006: Corine Land Class 2006, CCI-LC: Climate Change Initiative Land Cover)

Experiment	Topography and land use database
HRP1	ASTER and CLC2006
HRP2	SRTM and CCI-LC

4) PHYSIOGRAPHIC MODEL DATABASES

The terrestrial data sets for the WRF model are built using NCEP geographical data. These consist in global data sets for soil categories, land-use, terrain height, annual mean deep soil temperature, monthly vegetation fraction, monthly albedo, maximum snow albedo and slopes. The highest horizontal resolution available in WRF model is 30 arc-seconds, approximately 1 km at the equator, and it is called GTOPO30. WRF uses land-use categories from the US Geological Survey (USGS) 24-category data set with a highest resolution also of 1 km. When very high resolution is required or the region is very complex from a topographical point of view, the accuracy of these databases is not enough, and it is necessary to couple higher resolution databases. In this research, we have evaluated different topographical and land use database information, shown in Table 4.

5) NUDGING OPTIONS AND DATA ASSIMILATION

The last experiments realized have been focused on the use of three-dimensional variational data assimilation (3DVAR). The Gridpoint Statistical Interpolation (GSI) system developed by the Development Testbed Center (DTC) has been applied. This system is based on Spectral

TABLE 1. Physics options experiments.

Experiment	PBL	Surface layer	Cumulus	Shortwave radiation	Longwave radiation	Microphysics
INI	YSU	MM5	KF	Dudhia	RRTM	WMS3
MPH1	YSU	MM5	KF	Dudhia	RRTM	WDM6
MPH2	YSU	MM5	KF	Dudhia	RRTM	SBU-Lin
LWR1	YSU	MM5	KF	Dudhia	RRTMG	Best microphysics
LWR2	YSU	MM5	KF	Dudhia	FLG	Best microphysics
SWR1	YSU	MM5	KF	RRTMG	Best longwave	Best microphysics
SWR2	YSU	MM5	KF	FLG	Best longwave	Best microphysics
CMS1	YSU	MM5	MS KF	Best shortwave	Best longwave	Best microphysics
CMS2	YSU	MM5	Grell 3D	Best shortwave	Best longwave	Best microphysics
CMS3	YSU	MM5	New SAS	Best shortwave	Best longwave	Best microphysics
PBL1	MYJ	Eta sim	Best cumulus	Best shortwave	Best longwave	Best microphysics
PBL2	QNSE	QNSE	Best cumulus	Best shortwave	Best longwave	Best microphysics
PBL3	ACM2	MM5	Best cumulus	Best shortwave	Best longwave	Best microphysics
PBL4	MYNN2	MYNN	Best cumulus	Best shortwave	Best longwave	Best microphysics
PBL5	MYNN3	MYNN	Best cumulus	Best shortwave	Best longwave	Best microphysics
PBL6	UW	MM5	Best cumulus	Best shortwave	Best longwave	Best microphysics
PBL7	GBM	MM5	Best cumulus	Best shortwave	Best longwave	Best microphysics
PBL8	Shin-Hong	MM5	Best cumulus	Best shortwave	Best longwave	Best microphysics

TABLE 2. Dynamic options experiments.

Experiment	Turbulence and mixing	Eddy coefficient	Horizontal diffusion 6th order	Horizontal diffusion 6th order factor	Damping	Damping coefficient
INI	Dif. 2th	Smagorinsky	No	0.12	No	-
DIN1	Dif. 2th	Smagorinsky	Knievel	0.12	No	-
DIN2	Dif. 2th	Smagorinsky	Knievel	0.36 (d03)	No	-
DIN3	Dif. 2th	Smagorinsky	No	0.12	Rayleigh	0.2
DIN4	Dif. 2th	Smagorinsky	Knievel	0.36 (d03)	Rayleigh	0.2

Statistical Interpolation (SSI) and it is prepared to be coupled into WRF. Our methodology considers nudging the model towards observations (defined as observational or station nudging) and analysis (defined as analysis or grid nudging). Grid nudging is recommended for coarse resolution, while observation nudging is recommended for fine scale, and both of them can be combined. In this contribution, we have tested both nudging options applied over different domains and considering meteorological observations from: metars, radiosoundings and monitoring stations (data defined as global data); irradiance information from EUMETSAT satellites (data defined as satellite data); and from local meteorological stations managed by AEMET (Spanish National Meteorology Agency). This information has been coupled and combined into WRF defining different testing experiments as we show in Table 5.

3. Results and conclusions

The evaluation performed is focused on the inner domains, d03 and d04, since the final aim of this study is to find the best model setup for these areas. Statistical evaluation of the meteorological data is achieved by comparing the modeled parameters to the meteorological station observations of temperature at 2 m, wind speed at 10 m, wind direction at 10 m and relative humidity at 2 m. Wind speed and wind direction are calculated considering calms below 1 m/s, as wind direction is not reliable for lower speeds. The statistics have been calculated from hourly data of the model and observations.

Four statistics have been selected: Mean Bias (MB), Mean Absolute Gross Error (MAGE), Root-Mean-Square Error (RMSE), Index of Agreement (IOA) and Directional Accuracy (DACC). All results are shown in Table 6.

TABLE 5. Nudging options experiments.

Experiment	Nudging options d01 (9 km)	d02 (3 km)	d03 (1 km)
INI	NO	NO	NO
NUD1	Grid (global data)	NO	NO
NUD2	Grid (satellite data)	NO	NO
NUD3	Grid (global and satellite data)	NO	NO
NUD4	Grid (satellite data)	Observations	NO
NUD5	Grid (satellite data)	Observations	Observations

TABLE 8. Best model configuration selected.

Scheme or parameterization	Option selected
Initialization	GFS 0.25°
Microphysics	SBU-Lin
Longwave radiation	RRTMG
Shortwave radiation	Dudhia
Cumulus	Kain-Fritsch
Surface Layer	MM5 similarity
Planetary Boundary Layer	YSU (d01,d02,d03)
	LES (d04)
Vertical levels number	36
Diffusion 6th order option	Knievel
Diffusion 6th order factor	0.36 (d03)
Damping	Rayleigh
Topography	GTOPO30 (d01 and d02)
	ASTER (d03 and d04)
Land Use	GLC (d01 and d02)
	CLC2006 (d03 and d04)
Nudging	Grid (d01)
	Observations (d02 and d03)

In Table 7, the statistical evaluation of the different data assimilation numerical experiments is shown. In this case, the analysis period corresponds to an operational period of two months. Moreover, initial and boundary conditions are provided by GFS, while for the sensitive analysis CFSv2 is used.

As a result of this calibration, the WRF configuration that minimizes the uncertainty of forecasts for operational purposes in the region of Huelva is shown in Table 8.

A validation done for the period compressed between 01/01/2012 and 12/31/2013 using the selected configuration shows a confidence level of 70% for the temperature, 81% and 66% for the wind speed and wind direction respectively, and 90% for the relative humidity.

Acknowledgments. The authors gratefully acknowledge Port of Huelva and GTD System and Software Engineering for their support and collaboration.

References

Denby, B., and Coauthors, 2008: Guidance on the use of models for the european air quality directive. *ETC/ACC Report*.

Dudhia, J., and J. Wang, 2014: Wrf advanced usage and best practices. *16th WRF Annual Workshop*.

Garreud, R. D., and J. Rutllant, 2003: Coastal lows along subtropical west coast of south america: Numerical simulation of a typical case. *Monthly Weather Review*, **131**, 891–908, doi:10.1175/1520-0493(2003)131⟨0891:CLATSW⟩2.0.CO;2.

Krieger, J. R., J. Zhang, D. E. Atkinson, M. D. Shiski, and X. Zhang, 2009: Sensitivity of wrf model forecasts to different physical parameterizations in the beaufort sea region. *8th Conference on Coastal Atmospheric and Oceanic Predictions and Processes, San Diego*.

R. Arasa, M. O., M. R. Soler, 2012: Evaluating the performance of a regional-scale photochemical modelling system: Part i. ozone predictions. *ISRN Meteorology*, doi:10.5402/2012/860234.

Seaman, N., B. Gaudet, J. Zielonka, and D. Stauffer, 2009: Sensitivity of vertical structure in the stable boundary layer to variations of the wrf model's mellor-yamada-janjic turbulence scheme. *10th WRF Users' Workshop*.

Smagorinsky, J., 1963: General circulation experiments with the primitive equations. *Monthly Weather Review*, **91**, 99–164, doi:10.1175/1520-0493(1963)091⟨0099:GCEWTP⟩2.3.CO;2.

Stensrud, D. J., 2007: Parameterization schemes: Keys to understanding numerical weather prediction models. *Cambridge University Press*, doi:10.1017/CBO9780511812590.

Wang, W., 2014: Considerations for designing an numerical experiment. *16th WRF Annual Workshop*.

Warner, T., 2011: Quality assurance in atmospheric modeling. *Bulletin of the American Meteorological Society*, **92**, 1601–1610, doi:10.1175/BAMS-D-11-00054.1.

TABLE 6. Statistical evaluation for the months of February, May, August and October 2013.

Exp	Temperature (2 m)			Wind speed (10 m)			Wind direction (10 m)			Relative humidity (2 m)		
	MB	MAGE	IOA	MB	MAGE	IOA	MB	MAGE	IOA	MB	MAGE	IOA
	$< \pm 0.5\,^{\circ}$C	$< 2^{\circ}C$	≥ 0.8	$< \pm 0.5$ m/s	< 2 m/s	≥ 0.6	$< \pm 10^{\circ}$	$< 30^{\circ}C$	%	$< \pm 0.5\%$	$< 20\%$	≥ 0.6
INI	-0.40	1.82	0.97	1.30	2.24	0.64	10.97	31.41	68.61	-2.93	10.41	0.85
VER1	-0.51	1.77	0.97	1.37	2.19	0.64	9.90	33.47	69.36	-1.35	9.89	0.87
VER2	-0.51	1.73	0.97	1.37	2.20	0.64	10.10	33.43	69.27	-0.87	9.62	0.87
MPH1	-0.30	1.79	0.97	1.36	2.28	0.63	11.50	31.28	68.92	-3.01	10.50	0.85
MPH2	-0.21	1.78	0.97	1.39	2.30	0.63	10.61	31.06	69.54	-3.62	10.60	0.85
LWR1	-0.27	1.79	0.97	1.35	2.26	0.64	10.18	30.87	69.43	-3.61	10.63	0.84
LWR2	-0.31	1.80	0.97	1.35	2.27	0.63	10.99	31.75	68.25	-3.44	10.53	0.85
SWR1	0.02	1.77	0.97	1.38	2.26	0.64	10.21	31.57	68.45	-4.11	10.78	0.84
SWR2	-6.13	6.47	0.69	1.28	2.50	0.52	18.45	70.62	38.33	8.79	15.68	0.60
CMS1	-0.25	1.79	0.97	1.28	2.22	0.64	10.56	31.41	68.67	-3.47	10.71	0.84
CMS2	-0.27	1.79	0.97	1.31	2.22	0.64	10.49	31.20	68.88	-3.03	10.44	0.85
CMS3	-0.17	1.78	0.97	1.28	2.20	0.64	10.73	33.23	69.34	-4.49	11.07	0.83
PBL1	-0.04	1.79	0.97	1.68	2.43	0.61	10.02	32.95	70.34	1.06	10.79	0.84
PBL2	-0.32	1.91	0.96	1.63	2.39	0.61	9.21	33.17	69.64	-2.02	10.48	0.84
PBL3	-0.18	1.75	0.97	1.27	2.20	0.64	11.85	33.72	68.73	-4.98	10.95	0.84
PBL4	-0.78	1.88	0.97	1.37	2.21	0.65	12.99	34.27	67.72	-2.11	10.32	0.85
PBL5	-0.81	1.89	0.97	1.34	2.19	0.65	13.68	34.17	67.70	-1.85	10.28	0.85
PBL6	-0.29	1.76	0.97	1.28	2.22	0.64	10.90	32.99	70.36	-3.07	10.27	0.85
PBL7	-0.29	1.79	0.97	1.32	2.21	0.65	11.22	33.03	69.74	-2.93	10.34	0.85
PBL8	-0.21	1.78	0.97	1.32	2.24	0.64	10.82	33.38	69.46	-3.60	10.69	0.84
DIN1	-0.30	1.80	0.97	1.31	2.24	0.64	11.11	31.26	69.07	-3.17	10.48	0.85
DIN2	-0.29	1.79	0.97	1.30	2.23	0.64	10.99	31.19	69.19	-3.16	10.44	0.85
DIN3	0.37	1.81	0.97	1.30	2.25	0.64	10.96	31.39	68.84	-2.80	10.41	0.85
DIN4	0.30	1.79	0.97	1.29	2.23	0.64	10.47	31.17	69.23	-3.14	10.45	0.85
HRP1	-0.53	1.48	0.98	0.40	1.47	0.76	9.52	30.66	69.96	-4.44	9.45	0.89
HRP2	-0.41	1.55	0.98	0.90	1.88	0.68	9.89	30.87	69.48	-5.18	9.78	0.88

TABLE 7. Statistical evaluation for the period between 8/24/2015 and 10/24/2015.

Exp	Temperature (2 m)			Wind speed (10 m)			Wind direction (10 m)			Relative humidity (2 m)		
	MB	MAGE	IOA	MB	MAGE	IOA	MB	MAGE	IOA	MB	MAGE	IOA
	$< \pm 0.5\,^{\circ}$C	$< 2^{\circ}C$	≥ 0.8	$< \pm 0.5$ m/s	< 2 m/s	≥ 0.6	$< \pm 10^{\circ}$	$< 30^{\circ}C$	%	$< \pm 0.5\%$	$< 20\%$	≥ 0.6
INI	-0.57	1.78	0.89	1.13	2.02	0.67	4.19	39.65	57.63	-3.51	10.02	0.83
NUD1	-0.54	1.75	0.89	1.14	2.07	0.65	8.54	40.44	56.81	-4.53	10.50	0.82
NUD2	-0.54	1.76	0.89	1.13	2.04	0.65	6.07	40.28	57.83	-3.70	10.10	0.82
NUD3	-0.53	1.77	0.89	1.13	2.04	0.65	8.46	40.28	56.74	-4.55	10.49	0.82
NUD4	-0.55	1.77	0.89	1.13	2.05	0.65	7.39	39.28	58.57	-3.50	9.94	0.83
NUD5	-0.56	1.75	0.89	1.11	2.04	0.65	7.20	39.19	58.87	-3.42	9.91	0.83