Authors response to referees and interactive comments on the ACPD manuscript:

"Dry deposition of reactive nitrogen to European ecosystems: a comparison of inferential models across the NitroEurope network"

by C. R. Flechard et al.

We are grateful to the two anonymous referees for their positive comments and largely constructive criticism of the manuscript, which are answered in detail in the following sections. We have used a different font and colour (Arial blue) to distinguish comment from reply. Proposed additions to the original text are shown in *italics*.

Significant changes have been made to Section 3.3 (Comparison with flux monitoring datasets), which was split into 3.3.1 for NH3 and 3.3.2 for NOx, for which a substantial discussion has been added together with flux data from 3 sites (addition of Table 4). The end of the introduction was modified to state the objectives of the paper more clearly, and to specify and limit the scope of the study. A map of the inferential network will be added to the supplement, and is displayed on the last page of the present document.

We also thank X. Ge for the interactive comment concerning atmospheric amines, we have added their Atmospheric Environment review paper to the references.

****

1 – Reply to Anonymous Referee #1

General Comments

This is a well written and well organized paper. The model descriptions are extensive and a lot of material is covered in the paper. The approach of using the same site data for each model is informative. The paper contains a wealth of information. It is very informative to have so many sites included and to have them grouped by land cover type. While all major inorganic reactive nitrogen (Nr) species are covered, in the end the main focus of the model analysis, comparison and interpretation with observed data and commentary on research needs is on ammonia. While this is very valuable and important in the European context, it would have been valuable to spend some effort interpreting the model results for other major Nr species. With the inclusion of many sites covering four major land cover types and addressing annual Nr deposition this paper makes a valuable contribution to the literature.

The section on reducing uncertainties narrows down to only discussing NH3 and NHx fluxes. Are there no flux data for other species besides NH3? This points out a real need on the flux measurement side. It would be helpful to see something more added about HNO3, even the need for more extensive flux studies to reduced uncertainties (more than within canopy air chemistry). Hopefully the authors are considering more interpretation, analysis-related papers to mine the information presented here.

The "real need on the flux measurement side" is already highlighted in several places in the paper (introduction, Sections 3.3, 3.4 and in the conclusion, and in several of the referenced papers), but we have added a sentence in Section 3.4 on the need for more NH3 flux data (see specific comments below).

The criticism on the comparison of modelled vs measured dry deposition being focused on NH3 is a fair one. This results from NH3 being, together with NO2/NOx, the Nr species boasting the larger number of long-term flux monitoring datasets available in general, and within the NEU inferential network in particular. For HNO3, as mentioned in the manuscript p29321, l23, long-term flux datasets are non-existent within the NEU network. This is also the case for aerosol NH4+ and NO3-, with two exceptions being aerosol Nr flux measurements at NL-Spe and UK-Amo, but annual fluxes were not reported, their focus being on the understanding of variations in deposition velocity.

There are of course many other Nr flux datasets (including HNO3 and aerosol nitrate and ammonium) in the literature outside the NEU network, but a full-scale evaluation of the 4 models and comparisons with all available data is beyond the scope of this study, and has been done in part individually by each modeling team involved. We focus here on comparisons with existing datasets at NEU sites.

We thus propose to rename Section 3.3 with the following changes to the text:

"*Section 3.3 Comparison with micrometeorological flux monitoring datasets within the NITROEUROPE network*"

The introductory paragraph of Section 3.3 (p29321, l12 to p29322, l2) is modified thus:

"…for comparing scaled-up, annual estimates with actual, measured dry deposition. *In general, such long-term micrometeorological flux datasets are rare in the case of NH3 and NOx, and almost non-existent for HNO3 and aerosol NH4+ and NO3-. For the sites considered in this study, there are long-term data for NH3 and NOx only, at 5 and 3 sites, respectively, which are discussed in Sections 3.3.1 and 3.3.2; there are no available long-term datasets of HNO3 and aerosol fluxes. Aerosol deposition has been measured at NL-Spe (summarised in Ruijgrok et al., 1997) and UK-Amo (Nemitz et al. 2002), but annual fluxes were not estimated, the focus being on the understanding of variations in deposition velocity.*

In the few cases when long-term Nr flux estimates are available, the flux data capture is generally much lower than 100% and typically closer to 50% over one year; this means that measurement-based annual estimates are a combination of measurements and gap-filling and cannot be treated as absolutely accurate reference values, and are subject to some uncertainty. The procedures typically used in the annual datasets presented hereafter involved either the calculation of mean monthly diurnal cycles of measured fluxes, ensuring that season and time of day are properly weighted and accounted for; or the filling of gaps in the flux time series using inferential models with parameters fitted to local conditions (e.g. Flechard et al., 2010), or using neural networks (Neyrinck et al., 2007).

It should also be noted here that many forest sites of the NEU network have been monitoring wet-only or bulk deposition and throughfall as part of national or international initiatives (e.g. the ICPForests programme of the CLRTAP; http://www.icp-forests.org/), which, by difference between above- and below-canopy fluxes, have been used to provide estimates of dry deposition. However, uncertainties are large due to canopy interactions (Lovett and Lindberg, 1993; Zimmermann et al., 2006; Neirynck et al., 2007; Simpson et al., 2006b) and such data cannot be used reliably for model validation."

We then propose to split Section 3.3 into two sub-sections, *3.3.1 NH3* (starting p29322, l2, with the forest NH3 data), and *3.3.2 NOx*. A table (Table 4) is added and discussed in Section 3.3.2 for comparison of modelled NO2 deposition from this study with annual fluxes based on published flux measurements at BE-Bra, UK-Amo and NL-Spe. These are the only NEU sites with available annual NO2 and NOx budgets based on longer than annual flux monitoring data. NOx flux measurements were also carried out at a selected number of sites within the NEU project but the results are still being analysed and mostly unavailable (at the time of writing) for use in the present context. NOx data are discussed thus:

"…3.3.2 NOx

The only available annual NOx budget estimates based on long-term flux measurements are those at BE-Bra, NL-Spe and UK-Amo; NOx flux monitoring was also carried out at a number of other sites within the NEU project (e.g. CH-Oe1, FR-Gri, HU-Bug) but the results are still being analysed and unavailable at the time of writing.

The results for BE-Bra, NL-Spe and UK-Amo are summarised in Table 4. At UK-Amo, NOx flux monitoring has shown that NO2 dry deposition fluxes were small, in the range -1 to -5 ng NOx-N m-2 s-1, but also that the exchange was bi-directional with small NO2 emissions in summer daytime (Fowler et al., 1998). This results from NO emission by the underlying soil, with the oxidation by O3 to NO2 generating an effective compensation point for NO2 deposition; at low ambient NO2 concentrations, the ecosystem is a net source of NOx to the atmosphere (Pilegaard et al., 2001). In reality, it is at the soil level that a true compensation point exists for NO, which is driven by microbial nitrification processes close to the surface. Given the oligotrophic ecosystem and wet to water-logged peaty soil at UK-Amo, however, the soil NO emission potential is very low, so that the net annual NOx flux is downward and largely dominated by NO2 uptake by stomata. Bearing this in mind, the measurement-based NOx dry deposition estimate of -0.6 kg NOx-N ha-1 yr-1 is comparable with the model ensemble average of -0.7 kg NO2-N ha-1 yr-1 (Table 4).

By contrast, the annual measurement-based NOx budget for BE-Bra (above the forest canopy) is a net emission of +2.5 kg NOx-N ha-1 yr-1. This has been interpreted as the result of large NO emissions by the forest floor in this nitrogen-saturated Scots pine stand (Neyrinck et al., 2007), with the within-canopy oxidation by O3 of NO to NO2, resulting in a net apparent NO2 evolution from the stand. Downward NO2 fluxes were only observed at high ambient NO2 concentrations (>10-15 µg NO2-N m-3). High soil NO emissions, non-stationarity and chemical reactions in the air column between the soil, canopy and measurement tower in polluted environments hinder the interpretation of the total NOx flux (which is a conserved quantity) into its NO and NO2 parts (Fowler et al., 1998); thus no reliable NO2 dry deposition estimate could be derived for BE-Bra (Neyrinck et al., 2007).

Above canopy NOx flux monitoring at NL-Spe pointed to a net annual sink of -2.8 kg NOx-N ha-1 yr-1 for the year 1995 (Erisman et al., 1996). While this information also does not allow a direct comparison with modelled NO2 dry deposition from our study, and is subject to substantial uncertainty associated with the use of chemiluminescence and potential interferences by other NOy species, it can be set against available soil NO emission (dynamic chambers) data, which have been obtained at NL-Spe as part of several studies over the last two decades (Table 4). Early results from 1993 yielded an annual soil NO emission of +8.8 kg NO-N ha-1 yr-1 (Dorsey et al., 2004), but this was based on only a few days data in mid-summer. Later, as part of the NOFRETETE project and based on a more substantial dataset covering al seasons, Pilegaard et al. (2006) provided an annual estimate of +6.6 kg NO-N ha-1 yr-1 for 2002-2003. Unpublished results from the NEU project itself, and thus contemporaneous with our modelling study, indicate still lower annual soil NO emissions, of the order of +3 kg NO-N ha-1 yr-1 for the years 2008-2009 (Table 4) (A. Frumau, ECN, The Netherlands, personnal communication).

The comparison of net ecosystem NOx fluxes and soil NO emissions can only provide a likely range for the annual NO2 deposition from the atmosphere. Dorsey et al. (2004) showed that a large fraction (around 58% on average) of the emitted NO escaped out of the trunk space to react within and above the canopy at NL-Spe, but the fraction that was actually re-captured by foliage is unknown. Assuming a mean inter-annual soil NO emission of the order of +5 kg NO-N ha-1 yr-1, the maximum possible ecosystem NOx-N emission would thus be +2.9 kg NOx-N ha-1 yr-1, requiring a gross atmospheric NO2 dry deposition of (-2.8 - 2.9 =) - 5.7 kg NO2-N ha-1 yr-1 to yield the observed net NOx flux (1995 data). Conversely, if all the NO emitted from soil was recycled internally in the ecosystem, then the actual NO2 deposition from the atmosphere would only be -2.8 kg NO2-N ha-1 yr-1, which is in the range of values predicted by the inferential models.

These data illustrate the complex nature of NOx deposition, the inability of current inferential models to deal with bi-directional exchange, and the difficulty of finding long-term NO2 deposition datasets to validate models. Net NOx deposition only occurs at NO2 concentrations in excess of the canopy compensation point; this mechanism is only included, in a rudimentary manner (4 ppb

threshold), in the EMEP-03 model. A mechanistic treatment of this effect in inferential models requires the knowledge of the magnitude of soil NO emissions and of within-canopy chemistry and exchange. The prediction of soil NO emissions on the basis on N deposition and other environmental factors (Pilegaard et al., 2006) could provide a first step in the direction of an integrated ecosystem NOx exchange approach.

Table 4. Annual NOx exchange based on flux measurements at three NEU sites, and comparison with model results for NO2.

	Measurement-based annual NOx flux (kg N ha-1 yr-1)		*Modelled annual NO2 dry deposition* [†] *(kg N ha-1 yr-1)*
Site	*Above-canopy NOx flux (measurement years)*	*Soil NO emission (measurement years)*	*Average of 4 models (range) (2007-2008)*
BE-Bra	*+2.5 (1999-2001)* [¤]	*Not measured*	*-5.3 (-2.5 to -8.0)*
NL-Spe	*-2.8 (1995)* [§]	*+2.92 (2009)* [€] *+3.46 (2008)* [€] *+6.6 (2002-2003)* [¶] *+8.4 (1993)* [£]	*-3.2 (-1.5 to -5.2)*
UK-Amo	*-0.6 (1995)* [§]	*Not measured*	*-0.7 (0 to -1.1)*

[†] *This study*
[¤] *Neyrinck et al. (2007)*
[§] *Erisman et al. (1996)*
[€] *A. Frumau, pers. com. (2011)*
[¶] *Pilegaard et al. (2006)*
[£] *Dorsey et al. (2004); flux upscaled from only 3 days' measurements*

The following references have been added:

Pilegaard, K., Skiba, U., Ambus, P., Beier, C., Bruggemann, N.,Butterbach-Bahl, K., Dick, J., Dorsey, J., Duyzer, J., Gallagher, M., Gasche, R., Horvath, L., Kitzler, B., Leip, A., Pihlatie, M. K., Rosenkranz, P., Seufert, G., Vesala, T., Westrate, H., and Zechmeister-Boltenstern, S.: Factors controlling regional differences in forest soil emission of nitrogen oxides (NO and N2O), Biogeosciences, 3, 651–661, 2006, http://www.biogeosciences.net/3/651/2006/.

Pilegaard, K.: Air–soil exchange of NO, NO2, and O3 in forests, Water Air Soil Poll. Focus, 1, 79–88, 2001.

Fowler, D., Flechard, C., Skiba, U., Coyle, M. and Cape, J.N.: The atmospheric budget of oxidized nitrogen and its role in ozone formation and deposition. New Phytol., 139, 11–23, 1998.

Dorsey, J.R., Duyzer, J.H., Gallagher, M.W., Coe, H., Pilegaard, K., Weststrate, J.H., Jensen, N.O. and Walton, S.: Oxidized nitrogen and ozone interaction with forests. I: Experimental observations and analysis of exchange with Douglas fir, Q. J. Roy. Meteor. Soc., 130, 1941–1955, 2004.

Specific Comments

The title creates an expectation that some type of mapping exercise is to be undertaken to provide deposition estimates to ecosystems across Europe, which is not the case. A more accurate title might be: "Dry deposition of reactive nitrogen to four European ecosystem types: a comparison of inferential models across the NitroEurope network."

The title does not promise to deliver dry deposition estimates to all European ecosystems, or maps thereof. The idea, rather, is that a monitoring network such as the NEU inferential network covers enough vegetation types, climatic zones and pollution levels, to provide data coverage that can be thought of as being representative of European ecosystems in general. We believe this to be the case, and believe also that the suggested alternative title "Dry deposition of reactive nitrogen to four European ecosystem types" would be too restrictive.

Page 10, line 19: This gamma value seems awfully precise.

The gamma value is merely the ratio of the values of [NH4+]=600 µmol/l and [H+]=10^(-6.8) as used in the paper by Smith et al., 2000, and specified in the text p29303, l19-21. We have kept the same precision as given by Smith et al.

On page 17, use of this stomatal compensation point was deemed inappropriate when fertilizer was applied. Thus, this is for periods other than when fertilizer was applied, later termed "background conditions". This restriction should be noted here.

We have added the following text immediately following Eq. 11, just before section 2.1.3:
"*The canopy compensation point approach described here is applicable to crops and grasslands only outside periods of mineral or organic fertilisation, during which NH3 emission is governed by very different mechanisms (see Section 2.3.3).*"

Page 13, line 8: It would be very helpful to have a map of the sites somewhere, either in the main body or the supplemental material. This will help the non-EMEP reader interpret differences in conditions due to geography.

We agree that a map would be useful and have added one to the supplement, showing the measurement locations (this map is shown on the last page of the present document).

Page 15, line 7: It should be noted that the NO2 measurements from a chemiluminescence system will be biased high due to interferences from PAN and (unknown amounts of) HNO3.
We have inserted the following text on p 29309, l15:
"… Although NO2 concentration were not measured at all sites, *and although NO2 measurements by chemiluminescence systems are known to be biased high due to interferences by PAN and HNO3*, the available data are useful …

Page 15, line 31: While the NH4+ fine fraction may be higher than the SO4= fine fraction in the Netherlands (sea salt influence?) one would not expect this for more continental conditions. Are there no measurements elsewhere to be used?

We agree that, while the fine/coarse aerosol partitioning of NH4+ used in the MS is probably quite robust, the partitioning for SO4= is more variable and depends on the magnitude of the sea salt influence. There are in fact many factors affecting the fine fractions of ammonium and sulphate, not just seasalt sulphate, but also meteorological conditions and the relative abundances of ammonia and SO2 emissions.

Both ammonium and sulphate should be dominated by fine factions, and the fine fractions of these two particle species should be very similar under most conditions (including both inland and coastal locations). Although seasalt sulphate decreases the fine fraction of sulphate, it may not control the relative size of fine fractions between ammonium and sulphate. This is because ammonium and sulphate are closely related; if the fine fraction of sulphate is decreased by seasalt, the fine fraction for ammonium may also be decreased at the same.

As an example, below is a table (Table R1) showing fine fractions of the three particles species observed at eight Canadian rural sites (total 14 campaigns) (see Zhang et al., Atmos. Chem. Phys., 8, 7133–7151, 2008). KEJ1 and KEJ2 were two campaigns during two seasons at one coastal site, the rest were from inland sites. As can be seen, the fine fractions of ammonium were very close to those of sulphate, although slightly higher during a few campaigns. During one campaign at the coastal sites, fine fractions for both ammonium and sulphate were much lower than the rest of the campaigns, demonstrating the role of seasalt sulfate.

Table R1: Fine aerosol fractions (% of total fine + coarse), measured using Micro-Orifice Uniform Deposit Impactor (MOUDI) during fourteen short-term field campaigns at eight locations in both polluted and remote regions of eastern and central Canada (Zhang et al., Atmos. Chem. Phys., 8, 7133–7151, 2008)

	FRS1	FRS2	EGB1	KEJ1	KEJ2	ALG1	ALG2	LED1	LED2	CHA1	CHA2	SPR1	SPR2	BRL1
Pf(SO42-)	95	98	96	96	79	91	98	97	95	95	95	97	93	90
Pf(NO3-)	86	37	90	31	25	78	24	23	83	84	28	33	79	88
Pf(NH4+)	97	99	98	98	80	92	98	98	95	97	97	98	95	94

More relevant is the NO3- partitioning between the fine and coarse modes, which is also variable, as shown by Table R1. This does not only depend on the sea salt influence, but also on the probability of the sea salt (NaCl) to encounter nitric acid (HNO3), both of which are needed to form coarse NO3-. However, meteorology is actually the dominant factor, rather than seasalt, for the nitrate fine/coarse fractioning, due to the volatility of nitrate under hot weather. A detailed estimation of all these interactions across Europe is clearly well beyond the scope of this paper.

Since SO4= deposition is not actually reported in this paper (we focus on Nr dry deposition), this point is more or less irrelevant and we have therefore removed the reference to SO4= here in order not to confuse the reader. The text was modified thus, with the reference to *Zhang et al. (2008)* being added:

P29310, l9: …" *By default, and in a first approximation, fine aerosol was assumed to account for 94% of total NH4+ and 81% of total NO3-, following Ruijgrok et al. (1997), realising that in reality this ratio will be site specific, especially for NO3- (Zhang et al., 2008; Torseth et al., 2000), which has a larger contribution from coarse NaNO3 at coastal sites.*"

References added:

Zhang, L., Vet, R., Wiebe, A., Mihele, C., Sukloff, B., Chan, E., Moran, M.D. and Iqbal, S.: Characterization of the size-segregated water-soluble inorganic ions at eight Canadian rural sites, Atmos. Chem. Phys., 8, 7133–7151, 2008.

Torseth, K., Semb, A., Schaug, J., Hanssen, J. and Aamlid, D.: Processes affecting deposition of oxidised nitrogen and associated species in the coastal areas of Norway, Atmos. Environ., 34, 207-217, 2000.

Page 17, lines 12-14: A reference would be useful. One example is: D.R. Matt and T.P. Meyers, 1993. On the use of the inferential technique to estimate dry deposition of SO2, Atmospheric Environment, vol. 27A, No. 4, 493-501.

The reference suggested was added to the text.

Page 17, lines 23-25: Was this approach tested by taking a complete period and dropping out some hours/days to approximate how well this handled the issue? The tone of the sentence suggests not.

This was not actually tested, but since gaps in the availability of meteorological data were rare and randomly distributed, the procedure should not cause a large bias in annual scaled-up estimates.

Page 17, lines 29-30: Were the fertilization months excluded in the annual deposition total for all four models? The discussion suggests yes, but it would be helpful to explicitly state that here.

Yes, fertilisation months were removed for all models. This was made clear by modifying the sentence p 29312, l18-19 thus:
"…*Here the modelled (inferential) NH3 flux data from the fertilisation months were not included in the annual deposition total, for any of the four models, the reason being twofold; first, inferential models*…"

Page 18, lines 4-6: The point about inferential models over-estimating annual deposition where fertilizer application is involved is an important one. It would be useful to note here that this will be illustrated in Section 3.3.

We have added the following sentence at the end of the paragraph:
"...*The importance of field NH3 emissions by agricultural management events relative to background exchange is discussed in Section 3.3 by comparing model results with actual long-term flux datasets*".

This point will also be valid for vegetation having non-negligible compensation points. However, this point is nevermentioned again in the Reducing Uncertainties section or in the Conclusion section and it should be.

The point that a compensation point approach is needed for NH3 was made clearly in the paper, especially in Section 3.2, p 29320, l7-18, and in Section 3.4 ("Reducing uncertainties.."), p29325, l4-10.
That Rc-based models will overestimate annual dry deposition over vegetation with a significant compensation point is not a foregone conclusion; it very much depends on the non-stomatal resistance parameterisation used. For example, CDRY uses an Rc for NH3, but the Rext is also rather high (relative to the other models), so that annual NH3 dry deposition is the lowest of the four models.

Page 20, line 4: It looks like the IDEM model predicts larger values in Autumn as well.

We have modified the sentence p29315, l8 thus:
"…than the other models with typically 5–10mm s−1; *IDEM also tends to predict higher autumn Gs than the other models, especially for crops*. By contrast the EMEP-03 model …"

Page 23, line lines 27-30: Is it possible to provide the percentage of the dry Nr deposition from NO2, PAN and other NOy species at the CAPMoN sites so that we are comparing apples to apples? CDRY has a large fraction attributed to NO2.

The fraction of total Nr dry deposition attributed to NO2 by CDRY was large indeed. The following sentence was added to the manuscript p29319, l19:
"…of total (dry+wet) Nr deposition. *In that study, NO2 contributed 35% of Nr dry deposition, while PAN+PPN contributed 6%, NO 5%, HNO3 4%, aerosol NO3- 6%, other NOy species 11%, aerosol*

NH4+ 26% and NH3 just 7% (fractions averaged across the 14 sites). Most sites of the NEU network…"

Page 23, line 31: The 10-15% range seems only to apply to CBED, EMEP-03 and IDEM. Also, it appears that the top of the range is more like 20%.

The sentence was modified thus:
"…although NO2 concentrations were not measured everywhere, it may be assumed that NO2 generally contributed less than 10–20% of dry Nr deposition, as observed at e.g. IT-Col, FI-Lom, UK-AMo, HUBug*, despite the larger NO2 share predicted by CDRY* (Fig. 5)."

Page 24, line 7 and lines 11-12: Per page 17 comment, were periods of fertilizer omitted here? Please clarify for the reader.

Yes, this refers to background situations (outside fertiliser applications). The sentence was modified thus:
"…the compensation point approach in CBED allowed a few sites to be net annual emitters of NH3 and even of Nr (e.g. DE-Gri, IT-Cas) *in background conditions (without accounting for fertiliser- or grazing-induced emissions)*, while the other models…"

Page 24, lines 22-23: The inclusion of atmospheric N deposition in the list for emission potential (via the "or") seems like mixing apples and oranges. The compensation point response to N deposition is not the same as the emission potential, gamma.

The sentence was rephrased thus:
"…in the case of fertilised and managed agricultural systems, has long been recognized (Sutton et al., 1993; Fowler et al., 2009). New parameterisations for NH3 in CTMs are emerging (Zhang et al., 2010; Massad et al., 2010), which seek to relate the NH3 emission potential *to the plant/ecosystem N status, via total N inputs through atmospheric N deposition and fertilisation*."

Page 24, line 24: Suggest an insertion for clarity: ". . . atmospheric Nr inputs in excess of fertilization, since . . ."

The following change was made:
"…the challenge does not actually reside in the determination of atmospheric Nr inputs *in excess of fertilisation, since atmospheric deposition* represent typically less than 10% of added fertiliser…"

Page 25, line 5. Should note these fluxes are for summertime as well as daytime.

The text now reads:
"…PAN contributed about 20% of daytime, summertime Noy…"

Page 26, lines 3-8: There is not very good agreement between the throughfall and the flux measurements. This section illustrates the problem with using throughfall to provide a reasonable evaluation of the dry deposition estimates.

This point is also made on the previous page, p29321, l26 to p29322, l2.

Page 27, lines 1-14: It seems worth noting here that, given that there is likely some compensation point existing year round, then one would expect the models to predict high relative to the flux measurements, even for the background deposition. But they were mostly low. CBED is the exception. Only IDEM was high. A little more interpretation would be useful.

We do not agree that "…one would expect the models to predict [dry deposition] high relative to the flux measurements…" if the compensation point is ignored (as in Rc-models). At least, not systematically.

It mostly hinges on the treatment of non-stomatal deposition: one could expect Rc-based models to over-estimate NH3 dry deposition, relative to measured flux data in the case of a non-negligible compensation point, but only if stomatal and non-stomatal resistances were treated (modelled) correctly (compatible with reality). The paper shows clearly that model treatments of Gs (1/Rs) and Rext (or Rw, Rns) are very different, thus the models cannot be expected to show similar behaviours with respect to over- or under-estimation of dry deposition.

One Rc-based model might simulate an annual NH3 flux roughly compatible with the measured value, despite over-estimating stomatal deposition (ignoring the compensation point), by compensating with a non-stomatal resistance that is higher than reality. (This is only true for the net deposition case; a high Rext will not simulate a net emission, obviously). The model would thus get the numbers right for the wrong reasons.

Page 28, line 16: by "more extensive coverage" do you mean more varieties of landuse?

We mean more varieties of land-use and vegetation types, but also longer monitoring periods, not just monitoring campaigns of a few weeks in spring/summer centred on emission events. We also mean more pollution climates, with more measurements needed in arid and semi-arid areas, in the tropics, and in environments with high atmospheric NH3/acids ratios. The following text has been added:
"…which will not be bridged without a more extensive coverage of NH3 fluxes. *Specifically, more long-term (annual) datasets are needed for a wide variety of land uses, including broadleaf forests, crops and legume-rich grasslands, located in a wider range of pollution climates, with semi-arid and tropical regions a priority, and also in high NH3 environments, such as in the near vicinity of animal housing*. Within the NitroEurope IP,…"

Page 28, line 17: What species were measured at the intensive Nr flux measurement sites?

The Nr species investigated are specified in the modified sentence:
"… intensive Nr *(NH3, HNO3, NOx, NO3-, NH4+)* flux measurements to improve process understanding at a few core sites of the network…"

Page 28, lines 31-32: Good point.
Page 29, line 2: Would "inter-model differences" be better than "inter-model discrepancies"? Discrepancies suggests we know where truth is.

Online dictionaries provide the following definitions for discrepancy:
- "difference between two things that should be the same" (Cambridge)
- "an illogical or surprising lack of compatibility or similarity between two or more facts" (Oxford).
We believe these two definitions to be appropriate to the comparison of models in the present context.

Page 29, line 8: The statement "broadly comparable with the ensemble average" seems to refer to the ability to capture gross differences in deposition to different land cover types. There is a large inter-model variation at all of the flux sites and at 3 of the 5 sites the models seems to be biased low (when they should be biased high). A more nuanced conclusion would be helpful.

The paper does make it clear that inter-model variations are extremely large (See Figures 2 through 5), but our statement ("broadly comparable with the ensemble average") is there to suggest that, in

the context of inferential modeling in general, and not just this study, the use of several models and of an ensemble approach is one useful way of reducing the uncertainty in inferential deposition estimates.

Page 29, line 25: It seems that organic nitrogen should be noted. What about ". . . uncertainties in wet deposition estimates, including the lack of WSON, add to the . . ."

The suggested change has been made on p29326, l24:
"…uncertainties in wet deposition estimates, *including the lack of WSON data*, add to the total uncertainty…"

Technical corrections

Page 2 line 9: Insert "s" on "non-stomatal pathway are".
Corrected

Page 2 line 24: Suggest replacing "and" with "due to increased" to read "concentrations due to increased emissions by. . ."
Corrected

Page 22, line 19: Change "that" to "than" Page 22, line 23: Change "dry deposition to" to "dry deposition as"
Corrected

Page 25, line 8: At the end of the line should "known" be inserted? ". . . contribution of PAN and other known atmospheric organic nitrates . . .."
Corrected

Page 26, line 7: Should the second kg N be 18 kg NHx-N /ha/yr?
Corrected

Page 26, line 10 and 11: Should these be kg NH3-N /ha/yr?
Corrected

2 – Reply to Anonymous Referee #2

This manuscript presents estimates of the annual dry deposition fluxes of reactive nitrogen. The fluxes are calculated using the inferential technique, which combines measured concentrations and modelled deposition velocities. The concentration data are obtained from an extensive network of 55 sites across Europe. Four different dry deposition models are compared. The results of this study constitute an important contribution to the understanding of nitrogen balances and ecosystem loads. They also demonstrate that there are still large uncertainties involved in the quantification of nitrogen deposition. The paper is scientifically sound, clearly written and suitable for the scope of ACP. I recommend publishing it after the authors have properly addressed the following comments.

Major comments

(1) I would like to see a more coherent assessment of the main differences between the models. This kind of model comparison is stated as a key objective of the study (end of Introduction). The present version mostly points out differences in the results rather than explains their origin in terms of model formulations. Based on the current understanding of controlling processes, the plausibility of different parameterisations could be evaluated, rather than inherently assuming that all models are 'equal' (directly comparable) and thus the variation between the results could somehow be considered as an estimate of model uncertainty. In many respects, however, the models cannot be considered comparable. For example, in CDRY the parameterisation of the cuticular resistance of SO2 is directly adopted for NH3 without any justification from the measurement data (based on the scaling approach by Wesely who also did not have any NH3 flux data). This must be considered an ad hoc approach as compared to the actual parameterisations derived from an analysis of experimental data.

There exist several different approaches in the literature to validate dry deposition models. These include:

(1) Compare model results with field data. Many existing models were initially developed based on a small amount of field data and were also compared with field data at later times. However, a model might perform well at one location but badly at other locations with different biological and/or meteorological conditions. Thus, a large data set covering various scenarios is needed to fully validate a model.

(2) Model intercomparison identifying the 'origin' of the differences, i.e., the differences in model formulations, as suggested by this reviewer. For example, a recent model intercomparison study identified to what extent each resistance term contributed to the overall differences between the U.S. CASNeT and the Canadian CAPMoN dry deposition models (Schwede et al., 2011, Atmos. Environ. 45, 1337-1346). However, despite identifying the 'origin' in terms of model formulations, it is still not known which formula/parameterisation is more accurate than the others.

(3) Model intercomparison focused on their results. Because the end product of a dry deposition model is the dry flux it simulates, comparing the flux differences generated from these models will provide us a measure of the model uncertainties, although it cannot tell us how the model can be improved.

Approach (3) dominates in the present paper, even if a section (3.3) is devoted to comparisons with flux datasets (Approach 1), and if the comparative description of model frameworks and parameterisations (Sections 2.1 and 3.1) goes some way towards Approach (2). Since this study included five sets (gaseous and particle) dry deposition models, it is impossible to accommodate all the comparisons of each resistance term and/or their formulations. Instead, we only briefly discuss the major differences in their formulations (Section 2.1), and choose to focus on how large differences these models can cause on the results, i.e., the estimated dry N fluxes, at various

locations within the NitroEurope network. This paper is not meant to be a critical review of the different models. To make this clearer, the end of the introduction was modified thus:

P29296, l11: "…The primary objective of this paper is to provide an ensemble average estimate of Nr dry deposition for monitoring sites across the network, based on measured concentration data from the first two years of the project (2007–2008), and obtained by running four existing dry deposition schemes at the ecosystem scale.

A secondary objective of this study is to explore the differences in their output of modelled dry deposition and in their responses to input data, given the comprehensive dataset and wide range of vegetation types, meteorological conditions and pollution climates described by all monitoring sites. An alternative type of model intercomparison would focus on identifying the origin of the differences, i.e. the extent to which differences in model formulations and parameterisations contribute to the overall differences between dry deposition models (e.g. Schwede et al., 2011). Such an extensive analysis is beyond the scope of the present paper, however, as this study cannot accommodate all the comparisons of each resistance term and their formulations for four models and five major Nr species. Instead, the four routines are broadly described and compared with a view to point out the major similarities and differences in the approaches adopted by each model. We focus on the end products of the models, i.e. deposition velocities and fluxes (Section 3.1), the differences in which can be viewed as measures of current uncertainties in dry deposition estimates from inferential networks. In addition, comparisons with long-term measured flux datasets (Section 3.3) also provide scope for identifying priority areas of potential improvements."

Additional reference:
Schwede, D., Zhang, L., Vet, R. and Lear, G.: An intercomparison of the deposition models used in the CASTNET and CAPMoN networks, Atmos. Environ. 45, 1337-1346, 2011.

(2) The present discussion of the model differences seems somewhat illogical in that on one hand the authors argue that the models have been developed based on diverse data, resulting in large variation in results (p.29296, 3-5; p.29300, 23-27; p.29316, 25-26), but on the other they explain that there are very few datasets for model development (p. 29316, 19-21). As the datasets indeed are few, one would assume that the same data have been used for the development of different models and thus the predictions would tend to converge rather than diverge.

This point is not illogical when viewed from an historical and geographical perspective. As already mentioned in our reply to comment (1) above, a model can perform well at one location, but badly in other locations with different biological, chemical and meteorological conditions, if specific parameterisations are heavily empirical and reliant on observations obtained from a very limited number of sites, or even only one site. Indeed historically, the existing models were developed using different data sets collected in different countries or continents, over different vegetation or land use types, and under different climatic or meteorological conditions. These developments took place over time as our understanding of processes controlling surface exchange grew, but not necessarily in a concerted manner, not necessarily at the same time, and thus not necessarily using all the data that have become available over the years.

In fact, the very complex and varied responses of ecosystems as receptors (or sources) of atmospheric pollutants, as identified from the few available datasets with very different deposition velocities, resistances and compensation point, make it near-impossible to combine and reconcile all knowledge into a unified, coherent and fully mechanistic theory, even now with the benefits of hindsight and accumulated knowledge. Thus it is not surprising to see large differences when the models are applied for different scenarios.

The reviewer rightly points to various places in the manuscript where all these arguments are already developed; it would be superfluous to stress the point much further. We have modified the text towards the end of the introduction thus:

P29295, l28: "…These models have been parameterised on the basis of measured field flux data, but specific exchange processes and pathways are still poorly understood and their parameterisations remain crude and largely empirical. Also, model development has taken place in *different countries or continents, with different land uses, atmospheric chemistry, climates*, so that parameterisations derived from field data may not be universally valid. *Model development and validation tended originally to happen in parallel and be selective (rather than inclusive) in the flux datasets that were used in support. This was partly due to the very complex and varied responses of ecosystems as receptors (or sources) of atmospheric pollutants, observed in the few available datasets, which could not easily be reconciled and combined into a unified, coherent and fully mechanistic theory. This explains to some extent the very different existing parameterisations. With the increasing, though still limited, availability of Nr flux datasets, the knowledge and mechanistic understanding of surface -atmosphere exchanges grew over time, leading to increasing model complexity (big-leaf to multi-layer; dry deposition to bi-directional; fixed resistances to process-oriented).* Still, much variation in dry deposition estimates may be expected between models, hinting that uncertainties remain rather large."

(3) The paper would benefit from a more complete and unified description of the deposition models. This could be included in the Supplementary material. Just listing the different resistance terms with notation varying between the models (p.29300, 11) and referring to variables not defined in the present paper (e.g. p.29301, 5) is not very helpful to the reader.

The major differences in model formulations were discussed in Section 2.1. Since the focus of the paper is to discuss the fluxes differences generated from these different models, rather than identifying which individual resistance terms caused these differences (see reply to comment 1 above), a detailed model description, which can easily be found from the original publications for each of these models, seems not necessary here. However, we do discuss all the major differences between these models (e.g., stomatal and cuticle resistances, dry and wet canopies, compensation points) in Sections 2.1 and 3.1.

(4) As the focus of the study is on the annual deposition fluxes, long-term datasets have only been included in the model validation (Section 3.3). However, it is questionable if these datasets can actually be considered long-term measurements. According to Erisman et al. (1996), the data coverage of the NH3 flux data at NL-Spe is as low as 17%. At other sites the coverage may be better, but still is far from perfect. This means that, in order to obtain annual estimates, the measurement data must be complemented by inferential modelling, i.e. the technique used in the present study, but the available data indicate significant biases in the model predictions (Erisman et al., 1996). For a more realistic view of the data availability, the implications of data coverage and gap-filling should be discussed.

We agree that the data coverage is far from perfect. However, if data coverage is evenly distributed between the different seasons and the different times of day (that is, if the instrument down-time is randomly distributed), then the annual flux generated from these data should be a reasonable estimation, even with only 20% of the time. For NH3 flux monitoring, which in all 5 datasets of Fig. 6 was based on wet chemistry, the data capture was generally lower in winter due to frost and the practical difficulties of running the instruments in cold environments; night-time stability also reduces data coverage compared with unstable or neutral daytime conditions.

There are at least three ways of converting a discontinuous time series of NH3 or Nr fluxes into an annual value. The straighforward averaging of all available fluxes, scaled up to the full year, was not used in any of the NH3 data sets of Fig.6. The second way is to calculate mean measured daily diurnal cycles for each month, and to scale up this way. This should give proper weighting to the different seasons and times of day, and was done at NL-Spe and UK-Amo. The third way is to fill gaps in the data using an inferential model (CH-Oe1 and UK-Ebu) or a neural network (BE-Bra), though this still requires NH3 concentrations to be available when fluxes are not. A comparison of

the 2nd and 3rd approaches was given in Flechard et al. (Biogeosciences, 7, 537–556, 2010) for CH-Oe1, and showed that the third approach (using an inferential model) is best if model parameters (apoplastic NH4+/H+, stomatal and non-stomatal resistances) are fitted to the on-site measurements, which is akin to using a neural network.

We have added a few sentences to the introductory paragraph of Section 3.3:

P29321: "…*In the few cases when long-term Nr flux estimates are available, the flux data capture is generally much lower than 100% and typically closer to 50% over one year; this means that measurement-based annual estimates are a combination of measurements and gap-filling and cannot be treated as absolutely accurate reference values, and are subject to some uncertainty. The procedures typically used in the annual datasets presented hereafter involved either i) the calculation of mean monthly diurnal cycles of measured fluxes, ensuring that season and time of day are properly weighted and accounted for; or ii) the filling of gaps in the flux time series using inferential models with parameters fitted to local conditions (e.g. Flechard et al., 2010), or using neural networks (Neyrinck et al., 2007).*"

(5) The authors mention in the end of Section 3.4 that Nr flux measurements have been made at some sites of the NitroEurope network, but apparently these data have not been available for the present study. It might be useful to explain this in the introduction where NitroEurope is discussed, as the scarcity of validation data is a key limitation and the readers aware of this project may expect to see these new flux data to be used for model evaluation.

We have already adressed this point in our response to referee 1. In particular, we have added this sentence in Section 3.3:

"*The only available annual NOx budget estimates based on long-term flux measurements are those at BE-Bra, NL-Spe and UK-Amo; NOx flux monitoring was also carried out at a number of other sites within the NEU project (e.g. CH-Oe1, FR-Gri, HU-Bug) but the results are still being analysed and unavailable at the time of writing.*"

(6) In several places the chemical transport models (CTMs) are referred to, as the same parameterisations are used both in inferential flux calculations and CTMs. In particular, the model description (Section 2.1) starts with a discussion of CTMs. However, this discussion does not add anything significant and in some points may just appear confusing (and logically should be extended to equally relevant topics such as the treatment of the sub-grid scale). For example, it is obvious that modelled concentration data and grid-square-averaged meteorological data are used for CTMs, while inferential modelling is preferably based on directly measured input (Section 2.1). It is also obvious that 'surface' can be defined in many ways (Section A4), but this does not mean that the same flux-gradient profiles could not be applied within numerical weather prediction models (or subsequent off-line calculations) for physically consistent definitions.

CTMs are mentioned in several places in the paper in order to make the non-specialist reader aware that i) the models considered here are often used as part of larger modeling frameworks (the CTMs) and that the implications of the present study apply to some extent to the studies involving the regional models; and ii) that the results obtained here are NOT results of CTMs, which as we mention deal with grid squares of perhaps 50x50 km2, but local applications of the dry deposition routines. We believe it is important to make this distinction, before explaining in some detail how and why (e.g. p29297, l9, l19, l21; p29310, l6; p29313, l11; p29324, l24) model implementation differs here from the CTM version.

We have thus added a sentence p29297, l9, at the start of the paragraph, to clarify the differences between inferential vs CTM deposition model applications:

"*Note that the ecosystem/field-scale (inferential) application of dry deposition models, which is the topic here, should not be confused with regional (CTM) implementations of the same models.* For the CTM versions of the models, in which*…"*

There is also some incoherence in the related terminology (parameterisation, model, module, routine, SVAT) throughout the paper.

The various terms noted above have all been used in the literature before, with 'model', 'module', 'routine' and 'parameterisations' being used, sometimes inter-changeably, in the papers referred to in this manuscript.

(7) The implications of the sensitivity tests should be expressed in a more explicit way. The sensitivity to the definition of the nominal surface height is striking, but this also depends on the reference height of the data used for the parameterisations.

As noted p29318, l23, a full sensitivity analysis of the models is beyond the scope of this paper. The main objective is to show that the large uncertainties in dry deposition estimates are driven by the choice of input data as well as by model formulation. The text p29318, l23 was modified thus:

"…these results show that models have different sensitivities to input data and that the various land use classes respond differently. *These tests also demonstrate that uncertainties in inferential dry deposition estimates could be reduced by the on-site recording of vegetation parameters (LAI, hc). The uncertainty associated with surface potentials (T, RH) depends on the experimental conditions for the data on which the paramerisations were originally based. For non-stomatal resistances, ambient (e.g. 2 m above canopy) values have often been based, though not always (e.g. Flechard et al., 2010), while for the measurement of leaf stomatal conductances, temperature is measured in situ in a leaf cuvette."*

Specific comments

p.29296, 24-.: These references do not indicate that CBED and IDEM are "used within chemical transport models".

The following references have been added to the list and do refer to CTM applications of the models:

For the dry deposition model of CBED:
Vieno, M.: The use of an Atmospheric Chemistry-Transport Model (FRAME) over the UK and the development of its numerical and physical schemes. PhD thesis, University of Edinburgh, 2005.

For the dry deposition model of IDEM:
van Jaarsveld, J.A.: The Operational Priority Substances model. Description and validation of OPS-Pro 4.1, RIVM report 500045001/2004, RIVM, Bilthoven, the Netherlands.

Section 2.1.1: It would be more logical to start from Eq. (1).

We have shifted l19 – l24 (p29298) back to the start of the section, and have modified the text thus:

"2.1.1 Trace gases

The surface-vegetation-atmosphere transfer (SVAT) models use broadly similar resistance frameworks for pollutant trace gas exchange. In its simplest form the dry deposition flux F is given as the product of concentration at the reference height (zref) by the deposition velocity at the same level Vd(zref):

$$F = -(zref) \cdot Vd(zref) \quad (1)$$

with, by convention, negative fluxes denoting deposition, and Vd the inverse sum of resistances in series:

$$Vd(zref) = [Ra(zref,d+z0)+Rb+Rc]^{-1} \quad (2)$$

The atmospheric aerodynamic resistance, noted Ra…"

p.29298, 17; p.29307, 14: The figures should be numbered consecutively in accordance with their appearance in the text.

Figure A2 is actually a figure from the supplement and thus does not need to appear after Figure 1 of the main paper.

Eq. (3): This function is not used in all models. The light response term is different in EMEP-03, while CBED does not include the f_e and f_w terms.

The text has been modified thus (p29299, starting l16):

"…the effects of increasing vapour pressure deficit (vpd), plant water stress and temperature, respectively (Jarvis, 1976). *The fw factor is set to 1 in all models except in CDRY, where it is actually parameterised as a function of global radiation. The fe function is also set to 1 in CBED*. The last factor fs is a scaling factor to account for the difference in molecular diffusivity between water vapour and the trace gas considered. For the EMEP-03 model, *the light response term is different and* a further factor for phenology is also included (Emberson et al., 2001; Simpson et al., 2003)."

p.29299, 20, 25: If G_s =1/R_s is the bulk stomatal conductance (with LAI saturation), then R_s is not expressed on a unit leaf area basis.

The text has been modified thus:

"…in CBED, CDRY and EMEP-03, the bulk stomatal conductance Gs *(= inverse of bulk stomatal resistance*) does not increase linearly…"

p.29300, 15: In Eq. (3), R_s is defined for trace gases, not water vapour. The measurements of stomatal conductance are for water vapour, not trace gases.

Rs was replaced in the sentence by Gs.

p.29307, 11-13: It is difficult to understand why the model default LAI values would be better estimates than the direct measurements.

This is explained in detail in the supplement, Section A2. The direct measurements were not available at all sites nor for all seasons; they were often one-off measurements and their documentation did not distinguish between methods employed, one-sided, double-sided, or total LAI. Efforts to obtain this information from the network were unsucessful, which we regard as a great shame as LAI is such a basic parameter when interpreting micro-meteorology and deposition fluxes.

p.29314, 16: If G_s denotes the stomatal conductance of water vapour, then the notation should be changed, since R_s = 1/G_s includes the diffusivity scaling. Otherwise, the trace gas considered must be specified.

We have changed the phrase “Gs=1/Rs” to “*Gs (= inverse of bulk stomatal resistance to water vapour*)”

p.29321, 24-29; 9.29522, 16: It is not clear whether the authors consider the throughfall data reliable for the estimation of dry deposition fluxes. It is explained that lots of such data would be available that "may provide estimates of dry deposition, though uncertainties are large", but these data are not used. However, throughfall data from BE-Bra are used for evaluating the model performance at this site.

We have adressed this point in our response to Referee 1, by making the following change to the text p29322, l2:
“…which, by difference between above- and below-canopy fluxes*, have been used to provide estimates of dry deposition. However, uncertainties are large due to canopy interactions (Lovett and Lindberg, 1993; Zimmermann et al., 2006; Neirynck et al., 2007; Simpson et al., 2006b) and such data cannot be used reliably for model validation*.”

p.29322, 18: What is meant by ’total’ here?

The word ‘total’ was a mistake. This must read ‘measured dry deposition’

p.29322, 19: I cannot find any data for 1994.

Correct. The data are for 1995 only. This will be changed in the text.

Table 1: Why the ’measurement type’ is not specified? It makes little sense to compare one-sided and total LAI values.

This is true, of course, but, as written in the response to a previous comment, and explained in the supplement, Section A2, the database did not contain the information. Which is why we chose to use model default LAI for the base runs.

Fig 1c: It is not possible to seee the differences between the models for SN, G and C.

The data were pooled into one single category, short vegetation, as opposed to forests, since for u* large differences are only visible between forests and short vegetation.

Fig. 4: This figure is unnecessary. It is not discussed in the text and the same data are presented in a table.

We do not concur. Table 3 presents the speciated Nr deposition numbers, but not the total (all 5 gases + aerosols species added). Figure 4 is useful for showing graphically the large differences in total dry deposition between the models.

****

Additional corrections by the authors:

1: change Petrodolinskoye latitude from 42.5 to 46.5 in Table 1
2: Remove 2009 in “Zimmermann et al. (2006, 2009)” p 29294, l28.
3: p29293,l 16: “…also because NH3 inputs to fertilised, agricultural systems ARE limited by the presence…”

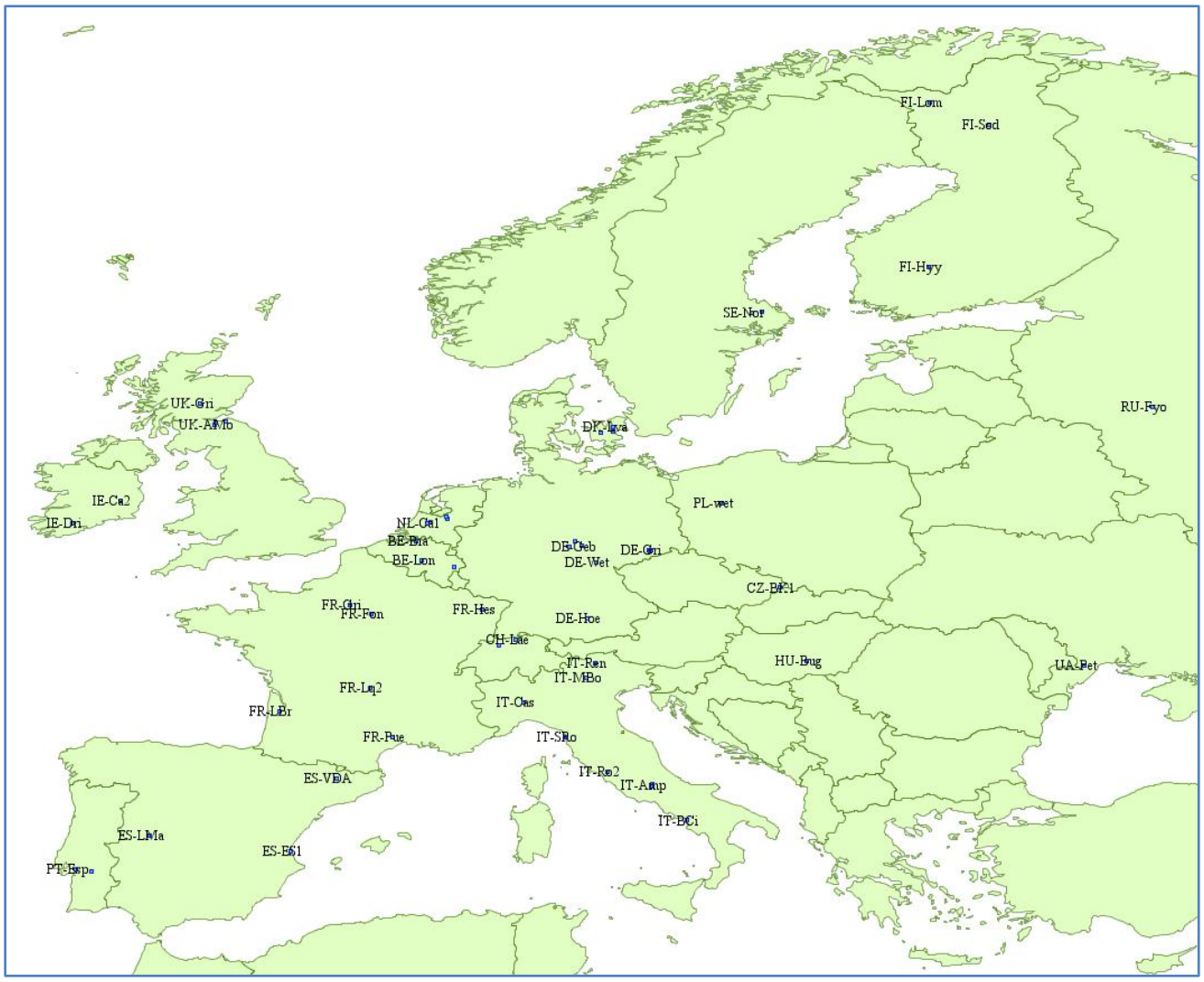

Figure A3. NitroEurope inferential network sampling locations.

Additional figure for supplement, showing NEU network sampling locations