Report on the PhD thesis by Jana Knappova

The thesis is an interesting and inclusive work assessing possible mechanisms regulating species diversity in abandoned fields and, more in general the distribution of grassland species in a landscape. The premise of the work is that these ecosystems provide the optimal study model to understand the importance of species establishment and dispersal and, in general, the dynamic of biodiversity in semi-natural grasslands. The thesis is composed by an introduction and 4 chapters in form of manuscripts (one being already published). The candidate showed a great capacity to combine observational studies with specific experimental manipulations to answer relevant ecological questions. Various modern methods of data analysis are applied properly and combined together efficiently. The results of the thesis are of great scientific relevance and are both novel and pushing the state of the knowledge a bit further.

The thesis is generally comprehensive and enjoyable. Nonetheless, the first more general comment is that it is somehow a pity that a final chapter summarizing the conclusions of the 4 manuscripts is lacking. This would have improved the relevance of this already interesting, work. It is true that the abstract and the "main results" section in the Introduction are already a remarkable synthesis of the work but, formally speaking, an extended conclusion is usually closing most of the PhD theses. In this sense the introduction of the work is including some results and discussion that could have moved, and enhanced, in a final conclusive section. Similarly, often the introduction is used to give clear theoretical background and explaining the general methodological approaches of the experimental setups considered. By giving much space to results, however, the introduction is, in this sense, a bit too much condensed. I understand that all main concepts and hypotheses have been presented properly. To me, however, an introduction of a PhD thesis should start from the assumption that the reader is not so familiar with the experiments being carried out, the study system, and concepts being introduced. To make an example species niche and species traits concepts are introduced as all readers were familiar with these concepts and applications. However, for example, no distinction between alpha and beta species niches, neither between niche position and breadth are given. No formal definition of species traits is given neither. These apparently trivial points, at the same time, seem to have consequences on the way concepts are handled in the 4 manuscripts. For example using Ellenberger values as traits is formally questionable, although the tests were surely informative. Another example is defining "availability of seeds in the surrounding landscape" as a trait (abstract Chapter 3) which is also questionable (although in the methods of this chapter authors better recognise their approach in reference to Violle et al. 2007). Having said so, I am sure the quality of the work is remarkable and the results and conclusions of this study are of primary importance and are very well documented. Only the feeling is that the envelop of the thesis could have been developed further. I am sure there were practical reasons to do so. Finally, another more general conclusion from the reading of the different manuscript is that, having time available, further analyses could be done to exploit even more the data. I hope the comments below will stimulate some further analyses for the papers not already published or submitted.

In the next section I will comment on each manuscript separately and raise comments and questions (**in bold**) that the PhD candidate might be willing to respond and discuss further.

The first manuscript, already being published, certainly deserves less comments as various referees have already raised most important comments. Therefore my comments are certainly one minor ones. The manuscript, overall, shows the importance of the quality of source habitats (i.e. number of species in this case) and the environmental characteristics of target fields in controlling the colonization and establishment of grassland species. I will surely cite this paper in the future because the results are interesting and clearly presented. It is shown that habitat characteristics of target fields are important for establishment. A simple question arises, **why habitat characteristics of the source habitats were also not taken into account?** A simple approach would have to compute environmental differences between target and source areas and use this as predictor together with spatial distance (and to perform a real partitioning of the dispersal vs. abiotic factor controlling the diversity of abandoned grasslands). Since environmental characteristics were estimated via GIS methods, I guess this could have been possible easily. I guess another simple indicator of the quality of target abandoned field could have been vegetation height and/or vegetation cover. In general the predictors in Table 2 for the abandoned fields could have been completed with more field measurements. Another point is the spatial structure of the data. **The description of abandoned grasslands**, right under Table 1 in the manuscript, left column, **might suggest that the 46 abandoned grasslands were, some, not spatially independent.** The figure in the Appendix seems to suggest this too. Data analysis: The data analysis was based on a linear regression approach where the number of target species was the dependent variable. I was a bit surprised to see this approach. I would have been interested to see differences in species composition rather than using the number of species as the only indicator. This would call, for me, to analyse the data in terms of species turnover between all plots (all abandoned and source grasslands) using multivariate analyses (e.g. see Chapter 4). For the discussion only, **I would be interested to know if the candidate thinks that considering multivariate analyses, i.e. focussing in species identity and composition instead of species number, would have changed the results, and in which sense.**

The second manuscript uses an approach of sowing 35 species (out of the 87 detected in the first study) from grassland habitats into 3 abandoned fields, in two years, and with and without neighbours. The results show that most of the species can establish. However the establishment is likely affected by both facilitation and competition of resident species and it is climatic and site dependent. The results highlight the importance of dispersal limitation and species interactions in determining species occurrence patterns in the landscape. I found the results of very high interest. The most interesting point to discuss, to me, is that I do not understand why the data analyses is mostly based on "only" 18 species. If the dependent variable is the number of established individuals (or flowering individuals etc..) I understand the problem raised by the authors. If the dependent variable would be success vs. non-successful establishment the whole 35 species could have been considered. **I think that comparing traits of which species made it (18) and which not (17 or 15), would also informative.** For example the effect of trait was assessed only on these 18 species, while a test showing differences between these 18 and the remaining 17 (or 15 if the two species that did not make it all are not considered) would have been interesting. **Are there differences between these two group of species?** Another question is comparing the species that flowered against those that established but did not: **are the traits of the species that flowered different from other species?** In this sense **it was also a bit difficult for me to understand fully the criteria why some species were**

included and some not in the analysis**.** If I did not get it wrong, **looking at table 1 and, e.g., species** *Aster amellus* **and** *Astragalus cicer* (the second being excluded but had rather comparable parameters and was even flowering) **reinforced this doubt.** A further minor comment is that **it is unclear what was the predictor when using traits.** I believe this was the mean of trait values, although it was not mentioned (or I have overlooked it). However even the **CV value for the trait values could have been interesting** (as showing intraspecific trait variability and if more "plastic" species adapt better than others). Other traits such as SLA, LDMC could have been of great interest (see comments on the next chapter). In summary I would have used traits not only directly into the model described at page 37 but also to test (1) which species established and which not, (2) directly on the data from Table 1, using separate columns as dependent variables (e.g. for flowering using flowered vs. non-flowered etc) (3) given the interactions with disturbance with species, assessing which were affected and which not by disturbance. **Overall I am not fully convinced the results reported at page 39, just above "spatial variability", and in the discussion ("species-specific reaction to disturbance" was not predicted by traits) can be interpreted in this sense.** As I understand the model run, described at the end of page 37, basically replied the species identity by species mean trait values. If I did not get lost, the dependent variable is not the species-specific response to disturbance (that would be, for example, establishment in disturbed minus establishment in undisturbed) but the establishment it self. **The final comment is regarding possible recommendations for the management of these abandoned fields.** Results indicate that bare soil is probably too stressed for establishment in more stressed climatic conditions (low precipitation years) but that flowering is decreased in undisturbed conditions. Maybe some more moderate disturbance, i.e. not bare ground but mowing, could be better for decreasing the competitive effect of resident. **Can mowing of abandoned fields be suggested to improve the establishment of species from the species pool? Results in this sense can be discussed even in light of the results from Chapter 4.**

In the third chapter an analyses is attempted to explain why some species are more, or less, abundant in the abandoned fields. The question being asked is the relative effect of availability of seeds in the surrounding landscape, dispersal traits and habitat requirements. Basically, the same database used for the first Chapter was considered. The results therefore first confirm, as in the first study, that the source of propagules is of high importance. In this sense it would have been interesting to **compare species that have a similar availability of propagules (i.e. present in many surrounding grasslands) but have different abundance in abandoned fields and compare their traits** (in this case maybe dispersal traits would show something). Could this comparison be possible? My second question is **why only dispersal traits were considered? Why also establishment and persistence traits were not considered?** These traits are supposedly also important in a way a species can really survive in a given site. Authors say that "frequency in abandoned fields (target habitats) as a measure of colonization ability" but it could be easily argued (also based on the other chapters of this thesis) that the frequency in the abandoned fields is also indicative of establishment and persistence abilities (including its competitive abilities). In this sense, because of these such tests were not considered, I am not sure I agree that species "quality" is not important in absolute terms. The conclusion of the paper is that both neutral and niche processes are important so that, anyway, authors recognize that species are not equivalent. Methodologically I am not sure I found explicitly said why in the Chapter 1 the 87 species were considered and **here "only" 32 species were**

considered, which were likely the better dispersed ones and most frequent. Of course this should give a higher importance to quantity than quality. How species were selected and was it possible to design other type of analyses to include also more species? Are there different traits between these 32 species and the 66 grassland species that all occur in the abandoned fields? I understand selecting species was required to have detailed information of seed availability, competition patterns etc. However the selection was probably based on the 32 most abundant species, therefore the test does not say much why some of the 87 species in the Chapter 1, and the 66 selected in Chapter 4, are more or less abundant. I would consider complementing this study using more "soft" traits from databases (LEDA etc.) for more species before safely concluding that quality is not so important. Finally, **Chapter 2 shows that germinability and competitiveness change across locations but here I guess that authors are using only a given value per species** (as I guess it was done, not sure how the values from 8 fields were condensed into a single value though).

The final chapter (IV) analyses colonization in interiors vs. edges parts of abandoned field, concluding that the edges are more easily colonized by grassland specialists. Authors conclude that edges are more easily to colonize rather because of less dense vegetation, suggesting better site conditions. However, chapter 2 suggests the less dense vegetation could prove more difficulties in emergence of seedlings but also, less competition from resident species. **As such I would like the candidate to re-discuss the patterns found in this final chapter with the results of the other works, particularly Chapter 2.** It is simply possible that **specialists that might prefer edges because they are more stress tolerant species and more poor competitors** (i.e. "Species with high light demands and low nutrient, moisture and soil reaction demands and species less frequent in dry grasslands were more restricted to field edges")? The description of traits in this chapter is mostly absent, at least in the methods, so that it is difficult to judge and again **other traits such as height, SLA and LDMC (even from databases) should be considered. Again this calls for confronting the traits of the 66 species that more frequent vs. less frequent in both interiors and edges and among them (interior vs. edges)** in order to complete the study.

In conclusion I found the thesis really stimulating and surely among the best that I red so far. I only suggest that few minor additional data analyses could further push the interest of the resulting manuscript. I wish the best for the promising career of the candidate.

With my best wishes,

Francesco de Bello