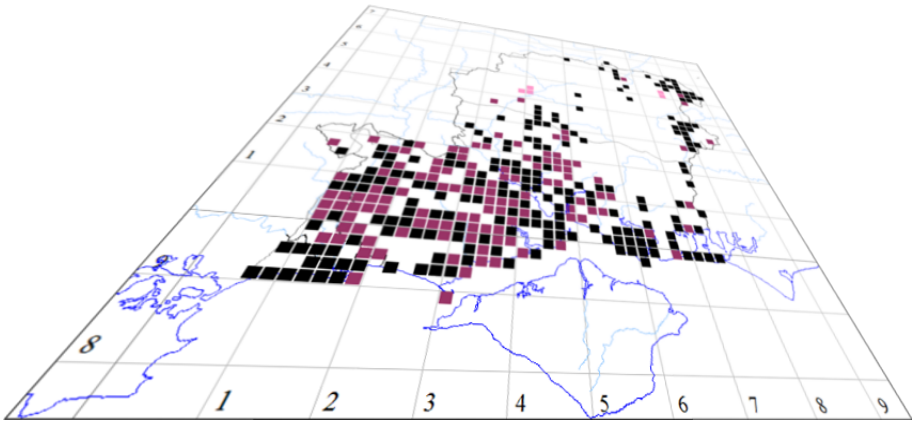
Chapter 5
Exchanging Data With Others



The author of this handbook is pleased to have errors pointed out and to receive suggestions for improvement and other comments.

Please send all communications to:
vc11recorder@hantsplants.net

CHAPTER 5 EXCHANGING DATA WITH OTHERS

Introduction

You need to read this chapter if you are going to share data with a community of other MapMate users (and that includes posting data to the BSBI hub), or if you are going to take records in electronic form from other people who are not MapMate users.

One of MapMate's strengths is its ability to allow data to be shared between MapMate users. If you followed the discussion on databases in Chapter 2, you will know that it achieves this without having any centralised repository of data to which all users are connected. Instead, you are one of a loose federation of recorders, each working independently, and you choose whom you will accept data from and who (if anyone) gets data from you.

The key to achieving this is the license code for your copy of MapMate, known to MapMate as the Centre Unique Key (or CUK for short). This forms part of the unique identifier that goes with each separate data record created with your licensed copy of MapMate, in MapMate-speak called the Global Unique Key (or GUK for short). It ensures that wherever your data appears in the big wide world of MapMate users, it remains recognisable as yours and it can't conflict with any other user's data.

When you exchange data with another MapMate user, using the standard facilities for doing so (known as 'synchronisation', or 'syncing' for short), MapMate doesn't just blindly take a dump of your data and put it onto their computer. That could lead to duplication of data if you didn't do it right, and it certainly wouldn't deal with deletions of records. Instead it goes through a rather clever process which:

- identifies all additions, changes and deletions that you might have made, and communicates these to your partners;
- ensures that if you do inadvertently send them the same data twice, it won't have any harmful effects;
- passes on any additions, changes and deletions that your other exchange partners may have sent you, so that you can act as a hub for disseminating as well as gathering information;
- keeps track of the most recent set of data that you have sent to each of your partners through a 'versioning' system, so that they do not get unnecessary repeats;
- uses the same versioning system to ensure that if you do end up getting a partner's information via two separate syncing routes, and the second set you receive is out of date, this is recognised and dealt with properly;
- allows you to control who gets what, through a filtering mechanism;
- prevents you editing other people's records once you have them, except by a special arrangement brokered by MapMate Ltd.;
- allows you to mark individual records so that they are not exchanged with other partners, either because they are confidential or unverified.

That's a pretty impressive list! In the next sections we describe the steps you need to go through to exchange data by this method, which are pretty straightforward for normal use. But then we take a brief look at how MapMate achieves these results, and the implications this has for the exchange process. This leads us to some “dos and don'ts” when deciding how to handle your fellow MapMaters, or trying to resolve problems that arise outside MapMate's control.

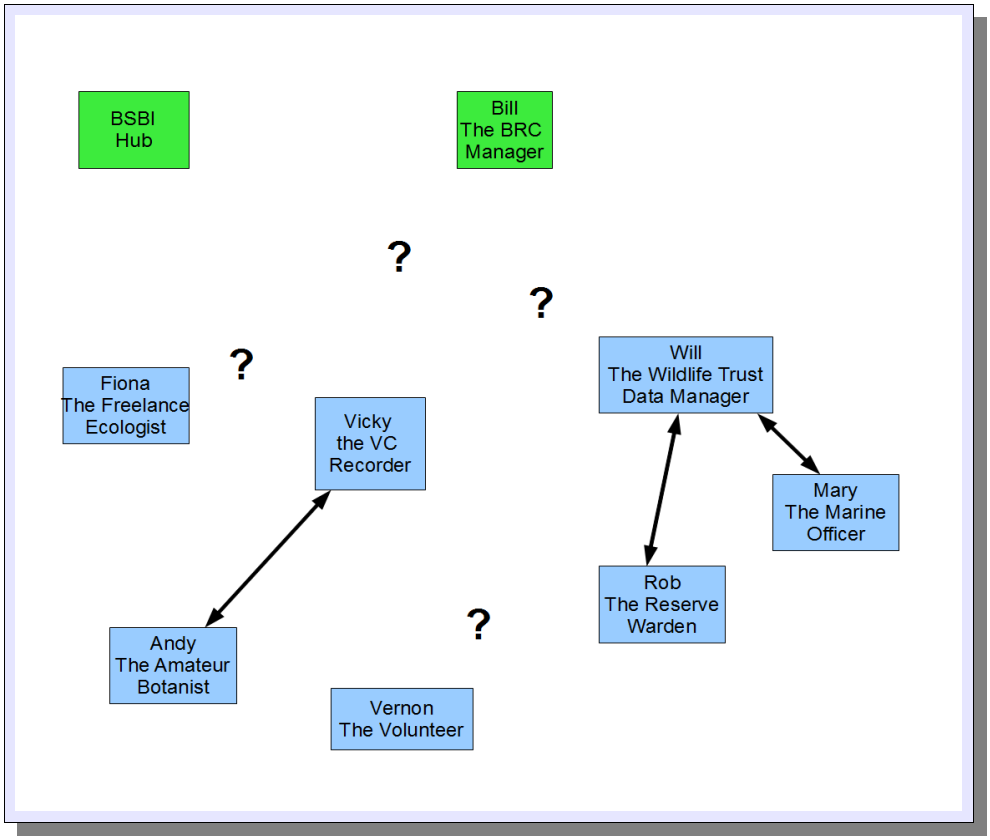
This is all well and good for getting data from other MapMate users who are actively maintaining their own record archive, but it won't deal with recorders who aren't using MapMate. They may well have kept their records on computer using a spreadsheet or another software package. Does this mean that you will have to re-enter all their data manually into MapMate? Thankfully, provided they can give you a minimum set of required data for each record, no. MapMate also has a mechanism for importing data from a standard format of text file that is easily generated from spreadsheets, databases and other software. It may mean that either they or you will have to do some rearrangement of their data, and supply default values for missing required items, but this is usually not hard and we describe how to do it later in this chapter.

Another situation you may be faced with occurs where a MapMate user is quite happy to send you his new records but doesn't want the burden of maintaining them thereafter. This can be handled through the same data export / import process as above, allowing you to take on his data maintenance, but requires a few caveats on how it should be operated.

We also look at the emergence of Web-based, map-based data recording facilities such as *LivingRecord*, and how these can be exploited in conjunction with MapMate to extend your constituency of recorders and improve the reliability of the information they send you. That leads us on naturally to means of validating the data you receive from others (and your own, before you send it on again!)

Organising Your Community of Data Sharers

Perhaps you will share data with only a few other MapMate users, and you will make exchanges directly with all of them. I am going to present a slightly more complicated scenario here, but the principles apply at all levels.



Put yourself in the place of Vicky the Vice-County Recorder. You have a community of recorders (and we'll assume they are all using MapMate) from whom you wish to acquire data, some or all of whom will receive data back from you. In addition you have local and national organisations who are also collating information from their associates at various levels.

There are many possible pathways that you can set up between these MapMate licensees. A few of them may be already established. Here I've assumed that the Wildlife Trust's officers will be sending data back directly to the Trust's HQ, and that you already have something going with Andy for records he submits anywhere in the vice-county.

The position is a bit less clear with the others. Vernon works on Rob's reserve, he also visits other Trust reserves and collects records, and he's an orchid fanatic who records them wherever he sees them. Fiona is quite happy that you receive her records but she already has an agreement to pass paper records to the BRC, and doesn't want to tie herself into any other formal agreements. You want to ensure that vice-county records find their way to the BSBI Hub, but you know that Bill the BRC Manager has already had some informal talks about this, too, and you haven't yet formalised an exchange agreement with the BRC.

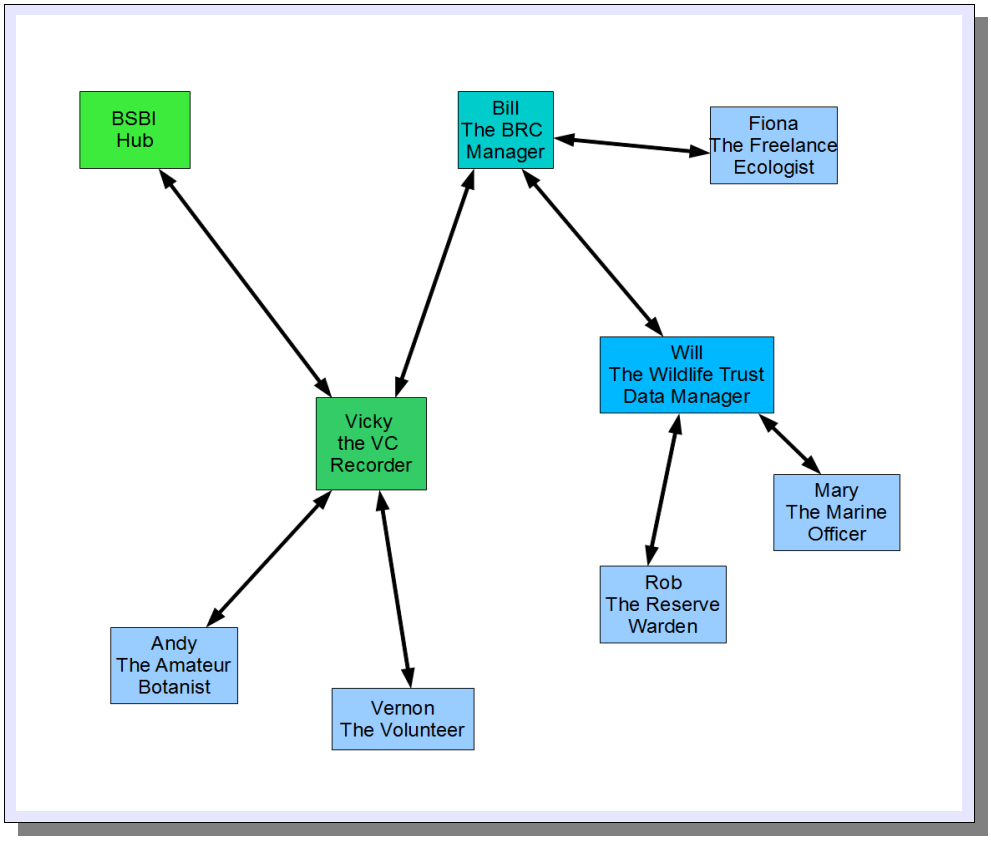
You need to talk to all the parties and come to an arrangement about how data flows

around this network. I propose one strong general "First Precept": **for any item of data, there should be a single path of exchange between any two MapMate licenses in the network**. In other words, it should be a "star" configuration with no loops; or in botanical terms, no anastamosing reticulations!

Some years ago I would have made this an immutable law, as there was an issue with MapMate re-using deleted records that could lead to data being corrupted when passed around a ring of users. This is fixed, and in general it seems that MapMate does cope well with data being passed in from multiple sources at different times. But I would still advocate it strongly, partly because it's better to be safe than sorry, and partly because such a scheme is easier to administer. If you are making two-way exchanges of data with at least some of your users, you will probably have a protocol for when you will pass data back to them. When you report on data in your system, perhaps for a published annual report or a book, you will be taking a snapshot of what you have available at the time. You don't want to be worried about how you stand with people who may be your partners some of the time, partners of partners some of the time, and perhaps having private exchanges with other people you exchange with the rest of the time!

A Strict "Star" Network

The simplest way to implement my First Precept is to bring it down to this, which is even more restrictive: **there should be a single path of exchange between any two MapMate licenses in the network**. Let's look at a possible scenario.



Because Vernon is making records from around the vice-county as well as Trust reserves, you agree with him and the Trust that he will submit all his records directly to you, and they will then come back to the Trust through the agreed channels for your interchange with them. But in the course of discussing that, you discover that the Trust and the BRC are already exchanging records. There seems no reason to disturb this arrangement, especially as they are using some of the same contract ecologists and can work out mutually beneficial arrangements for capturing survey data. So you agree that this will stay in place, and the BRC will act as the intermediary between you and the Trust for all record exchange.

However, you are aware that the BRC does not have the data validation and verification procedures that you would like to operate before passing data on to the BSBI national hub. So you diplomatically suggest that it would be best for you to deal with that exchange (giving you the opportunity to check data via the BRC as it comes in).

Meanwhile, Fiona agrees with the BRC that she will start submitting data via MapMate to the BRC, so that settles her position in the hierarchy.

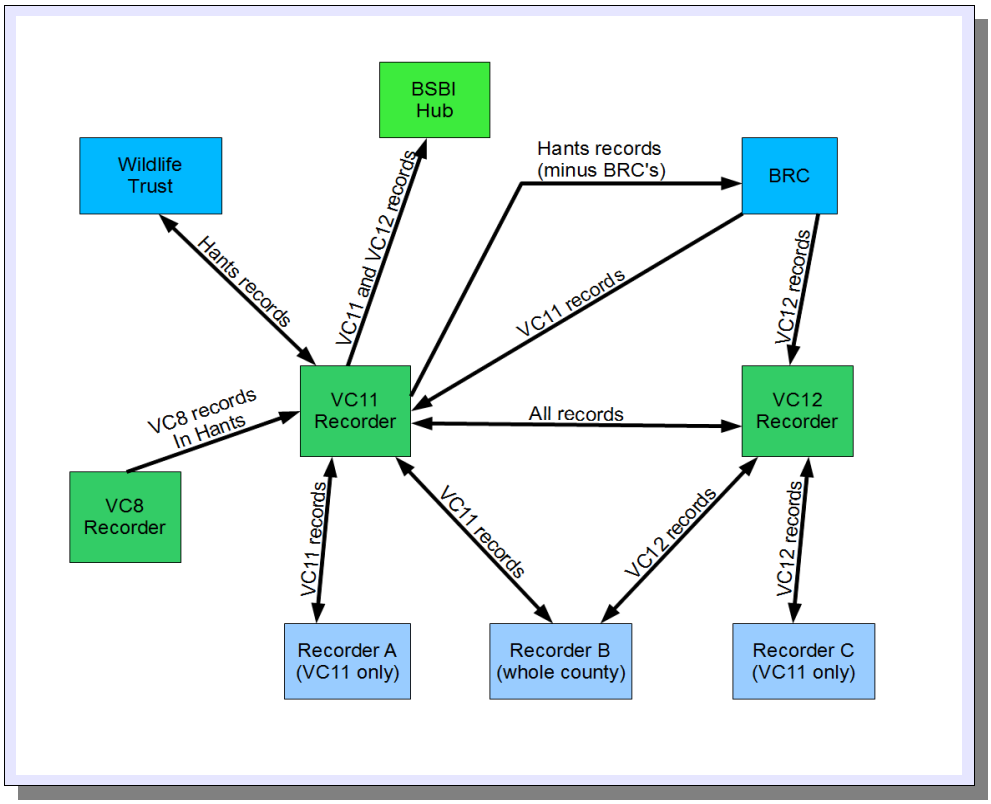
This gives us the network illustrated above, which as you will see gives us only one path for data between any two MapMate users. Under this arrangement, if we consider the BSBI as the main data hub, then Vicky is a sub-hub to the BSBI, Bill is a sub-hub to Vicky, and Will is a sub-hub to Bill. Of course it's not the only possible arrangement, and you will need to design your own configuration to suit circumstances.

Partitioning Data Exchanges

While I would always discourage sideways exchange of data between two MapMate users at a lower level in the hub hierarchy **where data may also be flowing up and down the chain**, there are sometimes circumstances where it is worthwhile to have the same person submitting different subsets of their data to different local hubs. As long as they are mutually exclusive subsets, this does not contravene the First Precept that **for any item of data, there should be a single path of exchange between any two MapMate licenses in the network**. I can best illustrate this by reference to my own situation. Mainland Hampshire is divided into two vice-counties; I am the BSBI recorder for South Hampshire (VC11) and Tony Mundell the recorder for North Hampshire (VC12). A sizeable chunk of VC11 is now in administrative Dorset, while there are areas of VC8 (South Wiltshire) and other vice-counties that are now in administrative Hampshire. We exchange data with the Hampshire Biodiversity Information Centre, our local BRC, which is concerned with administrative Hampshire; and the local Wildlife Trust, which is concerned with administrative Hampshire, and also the Isle of Wight which is yet another vice-county (but we'll leave that out of the picture here).

For convenience, it's very useful to have a collated set of records for all of mainland Hampshire (both the administrative county and the two vice-counties) that is under the control of the BSBI recorders. On the other hand, we each need to be able to play our separate roles in verifying data, and in some cases reporting on and analysing data, as vice-county recorders. To do this, we have a “binary star” arrangement with all data fully shared between us.

In order to operate such a system without confusion, we need to set up filters on the exchange process to say who gets what when an exchange takes place. I shall explain how this is done in the following section; for now, you just need to know that it can be done in principle. In the following diagram I've omitted a few pathways that don't relate to Hampshire / VC11 / VC12 records; it's complicated enough as it is.



At a casual glance, this looks as if it breaks my First Precept, but look at it from the point of view of any individual item of data from VC11, 12 or 8. There is only a single pathway by which such an item can flow from one part of the network to another.

Consider the case of Recorder B. He records across the vice-county boundaries, but by arrangement with the two recorders he has set up his MapMate exchange so that all his VC11 records will go to one and all his VC12 records to the other. This gives each recorder the chance to validate and verify the records for which he has responsibility.

At frequent intervals the two VC recorders exchange data with each other, so that each has a fairly up-to-date picture for the whole county – and if they need to be fully up-to-date, they just need to request a sync file from the other. If Recorder B gets other new data back he gets it separately from each of them. It didn't have to be that way; for instance, I could agree to be the 'master repository' and to send both VC11 and VC12 records back to him.

In fact, that's what happens in our relations with the national and county organisations. Updates to the BSBI Hub, the Wildlife Trust and the BRC tend to be on a less frequent, more or less scheduled basis. For each of these exchanges I make sure that I am up to date with Tony's data, and then feed the complete set on to the organisation. Because the BSBI works along vice-county lines, they get just the VC11 and VC12 data. The BRC and Wildlife Trust get just the administrative county data. The Wildlife Trust is a two-way exchange, so

when I get data back from them I will re-sync with Tony to bring him up to date.

The relationship with our BRC needs explaining. They don't actually keep their main record repository on MapMate, but on Recorder. Unfortunately there isn't a two-way data exchange between MapMate and Recorder that would take care of all the synchronisation issues, so this is how we have worked out our salvation locally. Fortunately, Recorder can export data in a form where it can be brought into MapMate. This is the Text Import format that we discuss in a later section. Tony and I each take records for our own vice-county and import them into our local own copy of MapMate. Because we are using the Text Import mechanism, they become "our" records, with our own Centre Unique Keys associated with them, giving us the right to amend or delete them. We then go through the process of checking them. Once we are satisfied, we perform an exchange of data between us.

That then leaves us with the problem of passing the data that we've acquired from other sources back to the BRC. For this purpose, they have a MapMate "posting database", so for this we can use MapMate's normal data exchange facilities. I do this exchange for both vice-counties' data. They then combine this with their own data into a third database linked directly into their Geographic Information System (GIS).

Because the records that we get from the BRC acquire a brand-new identity when we import them into MapMate, we have to be sure that we don't re-export them and saddle the BRC with the problem of duplicates. To this end, we run an automated process before importing the Text Import file that puts a text marker against each Record, saying that it has come from them. Our filter for exporting records to them uses this to block such records going back.

One-way Street or Two-way Street?

Probably, if you have exchange agreements with organisations such as those we've been discussing, those will be two-way exchanges. Beyond that, you may want to consider what's best in the case of individual recorders on your patch.

Clearly, getting feedback can be highly motivating for an individual, and if your more committed and prolific recorders are using MapMate then it makes sense to give them the benefits that MapMate bestows in this regard – especially if they have access to distribution maps through the Atlas feature (see Chapter 7). If they are seriously engaged in a collaborative recording project such as the BSBI National Atlas or recording for a county Flora, full two-way exchange is more or less a *sine qua non*.

Whether you want to extend this to all your MapMate community is worth considering on several counts. This brings me to my Second Precept: **be conservative in setting up two-way exchanges**.

In the first place, not all recorders may want data back from you. Especially if they are treating MapMate as their own personal archive, they may not want to be swamped by records for a whole county or vice-county.

Secondly, there is a question of principle. You need to ascertain that all your other providers are willing to have their data shared in this way. Most will happily subscribe to

contributing to national and local hubs for scientific and conservation purposes, although this is something that you need to make sure is clear to them. You certainly need to get their agreement to any more general sharing.

Finally, there is a practicality. As we've discussed in previous chapters, data acquired by data exchange from another MapMate user can't be edited by the acquirer; it remains in the ownership of the original supplier. But there's nothing to stop anyone **using** ancillary data such as Site information that came from a third person to support their own records.

Consider the following: User A submits a number of records to his local hub located at Site X, with a 6-figure grid reference. Subsequently the hub exchanges data with User B, and User A's data gets passed down to User B. When entering his own data at a later stage, User B enters a 6-figure grid reference into the Site box and finds there's already a site defined on MapMate for that reference.

Being a member of the Occam's Razor Fan Club, User B makes use of that Site reference and doesn't invent a new one of his own for his own record. He then submits that to the hub.

Then User A makes a new record for Site X, and in the process realises that he's mistyped the grid reference. So he moves it 300 metres to the east, and his new record (along with the amended Site information) goes to the hub on the next exchange.

But now, User B's record (which really was 300 metres west of User A's) has been invalidated on the hub, and there is nothing in any subsequent exchange process that will make that apparent to him – unless he's prepared to pore over every detail received in each sync he gets.

There are ways to manage this situation, but not to enforce an avoidance. For instance, you can ask that every partner sends you data based only on 'My Sites' (meaning his or her Site definitions). But unfortunately, there isn't a built-in Sites default for 'My Sites'. It can be implemented, but it requires a bit of technical fiddling and has to be done independently on the MapMate system of each partner. Even then, you can request that they use this default setting, but you can't enforce it. Since they would have to turn it on for data entry, and off whenever they want to look at all the data for an area, it's easy for them to leave it at the wrong setting.

Clearly the more people you have who can fall foul of this feature, the less tractable it is. So there are good grounds for making sure that there is a real need as well as a desire for a partner to receive other people's data back from you.

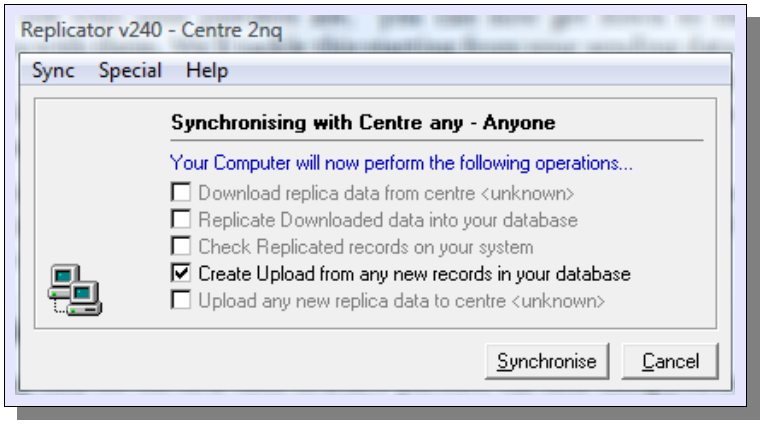
Running a Sync Exchange

Having sorted out who your partners are, you can now get down to the business of exchanging data with them. We'll tackle this starting from the position of your sending data on to another MapMate user. In reality, you may start off by receiving data from them. But if you read right through this section, you will get some of the background information you need for both sides of the process.

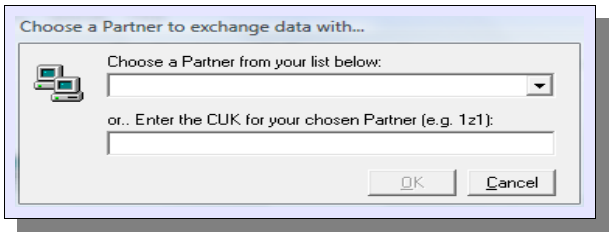
As we've mentioned earlier, MapMate keeps track of your additions, amendments and deletions, and it keeps separate track of what's already been sent to each of your partners in the past. You don't have to keep account of these things.

Sending Data

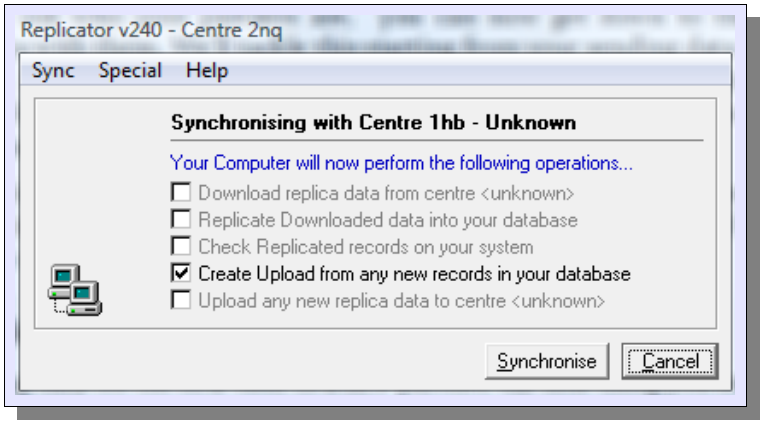
If you're sending data for the first time to a new partner, the first thing you need to know is their CUK. If they're not sure, they will be able to find out by invoking **Help / About MapMate** on their copy of MapMate.



Now select **Replication** from the main menu of the MapMate main screen. You will see a screen like this. Select **Sync / Change Default Partner...** from the menu of this form.



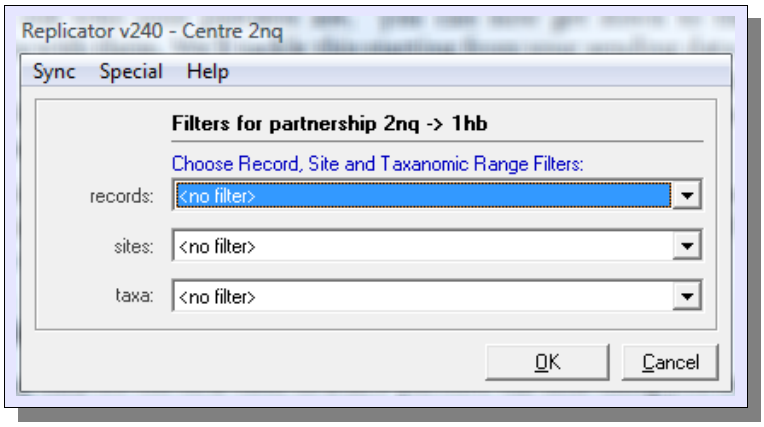
Now you will need to type the CUK into the bottom box of this form. Then click on **OK**. Once you've either sent or received data, that partner's code will appear in the drop-down selection list above.



You will now be back to the original Replicator screen, but it will look like this. This partner will continue to be listed as **Unknown**, both here and in the selection list on the last screen, until you receive data back from them. After that, their registered name will appear. If you only send to this partner, and don't receive data back, you might want to make a note

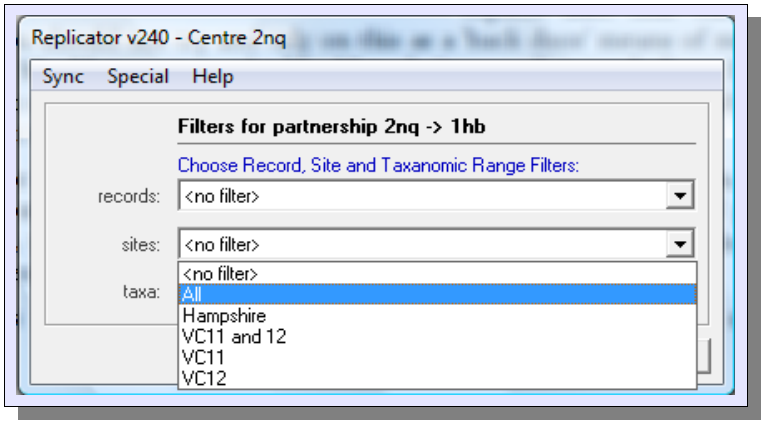
of their CUK for the future.

Don't click on the **Synchronise** button yet! The next step is to specify what data you will send them. Select **Special / Set Filters...** from the Replicator menu.



For **records**, the drop-down list you will see depends on which version of MapMate you have, and what system updates you've subsequently loaded. You will probably find some filters for specialist recording groups, and also some for records relating to specific recent years. Note that the latter refer to the date of the Record; MapMate doesn't keep track of the explicit date that the Record was entered. So you should not try and rely on this as a 'back door' means of managing your syncs to other MapMate partners; let MapMate do its own dirty work there. If you need to **export** data to other people by routes other than syncing, they can be handy – as long as you get all your records for a recording year entered up by a given date!

There are three basic options in the list: **<no filter>**, **All Records** and **My Records**. If you are an individual recorder using MapMate only to record your own records and perhaps view other people's, I recommend that you use **My Records**. If you are a hub or sub-hub, collecting data from others to pass on, you will need to use one of the other two. The difference is that **All Records** sends only the associated data (Sites etc.) that are linked to the records you are sending on that sync. So this is the most economical option for normal use. **<no filter>** will send associated data whether or not it goes with the records currently being sent. This can be useful if there are gaps in the data at the receiving end and your partner is not sure how extensive they are, but is otherwise probably best avoided.



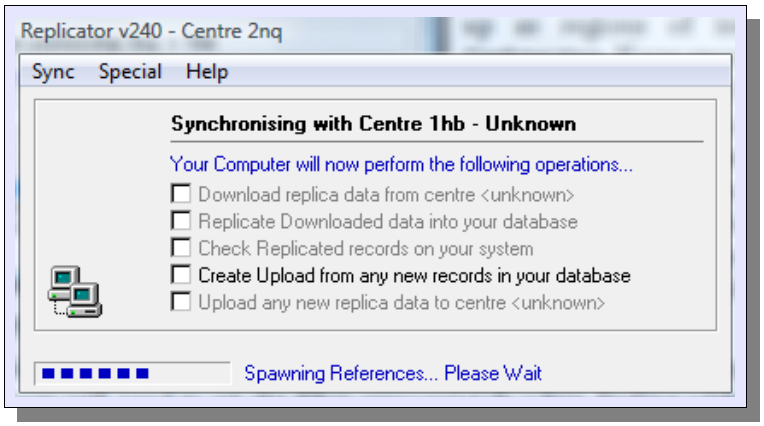
What appears in the **sites** filter list depends on what you previously set up as regions of interest in **My Configuration**. If you record in only one area, and your partner has an interest in all records for that area, then you can leave the setting at **<no filter>**. If you record in Dorset and Wiltshire as well as Hampshire, say, then you would want to set the filters for exchange with each vice-county recorder appropriately.

taxa is very similar. If you record birds, butterflies, dragonflies and bryophytes as well as vascular plants you will need to set the filter appropriately when dealing with a partner in a specialist group.

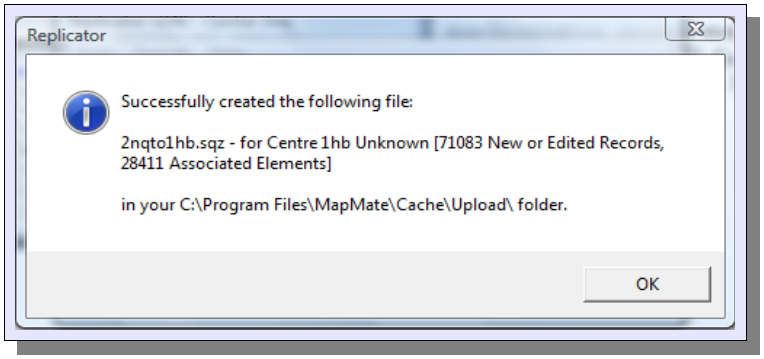
Having made your choices, click on **OK**.

Now you can click on the **Synchronise** button back on the main Replicator form. If this is your first sync with the partner you will see a prompt to confirm your filter settings. If you have changed the settings, you will see a further warning that you may need to resync **all** your records with this partner. For now, ignore this message; we'll discuss when and how you do that in a later section.

By the way, filters only apply to data you're sending out. They have no effect on anything you receive from your partners, so it's always the responsibility of the sender to get this right.

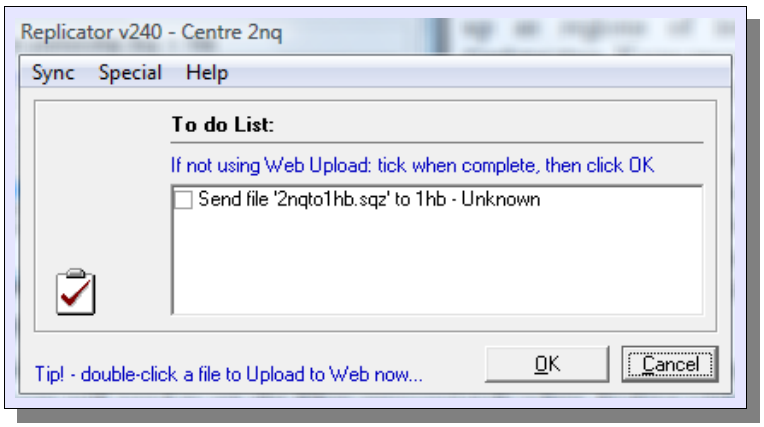


Depending on the number of records you've accumulated, the synchronisation process may take a little while. MapMate will display the stage it's at and a progress bar as it works.



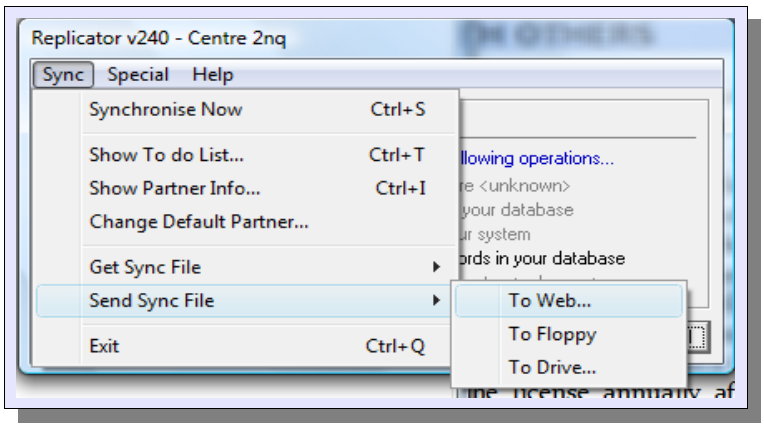
Once the process is finished, you should see a message like this. The 'Associated Elements' referred to are records of Sites, Recorders and other items that accompany your Records. Take note of where your exchange data has been stored; in my case, I ran this illustration on a MapMate license still using the 'old regime' locations for data. If you are up to date, you will probably see a location in the **Documents\My Mapmate** folder.

At this point, it's important to realise that getting your data to your partner is a two-stage process. So far we've only been through the first stage, which is to generate the exchange information into a file on your computer. You still have to get the data to its destination.



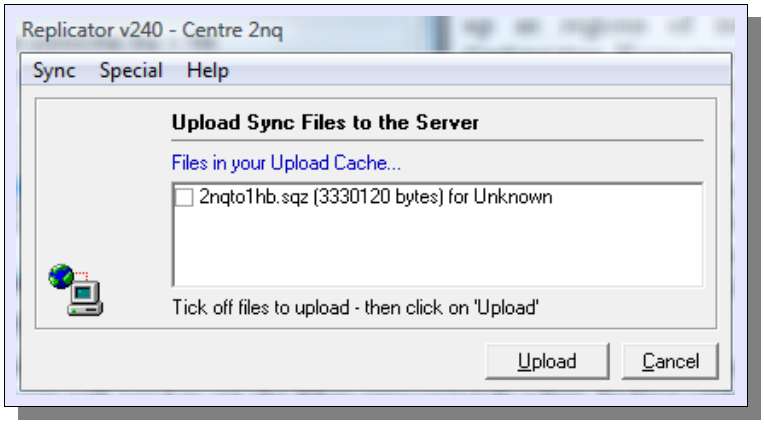
After clicking **OK** on the display shown above, you'll see a form like this. You might be tempted to think that if you tick the box beside the partner information and click on **OK**, that will take care of the transmission. It doesn't. This is just a checklist of things done and things to do. So for now, click on **Cancel**. But notice, just before you do, that

exchange files have a particular format of file name: “<your CUK>to<partner CUK>.sqz” in the case of files you're sending; “<partner CUK>to<your CUK>.sqz” for files you're receiving. There's no versioning or date-stamping there, so any existing file of the same name will be overwritten if a more recent one is generated or downloaded. It's important to make sure that they have been dealt with before you run your next sync.



You have a number of options for transmitting the file. From the Replicator menu bar, select **Sync / Send Sync File**. By far the most straightforward method is to select **To Web...**; to use this, you must obviously have an Internet connection, and you will also need to have registered your copy of MapMate with MapMate Ltd. (and renewed the license after your first year of use). With this method, MapMate Ltd. provide a staging post for your data on their own computer systems, and will alert the recipient that there is data available to download.

Sending to the Web



Once you select this method, you will see a list of files available to send. Tick those you want to send, and click on **Upload**. MapMate will then display a progress screen. When the transfer is complete, you will find that if you return to the To Do List screen (**Sync/Show To do List...**) the entry will be ticked off.

Things That Can Go Wrong While Sending

- **Your Internet connection is not available.** Remedy this, or wait for service to be restored by your service provider. By the way, if the connection breaks during the sending of the file, MapMate keeps track of progress and will resume from the point of failure the next time you send via the Web.
- **The MapMate Web service is unavailable.** This is usually a matter of waiting a while. If the problem persists, let MapMate Ltd. technical support know.
- **You have a firewall in place.** A firewall controls what programs are allowed access to the Internet, potentially exposing your computer to being seen by other systems on the Internet. This will not be a problem with reputable commercial or 'open

source' software, including MapMate, which only get access to your data for specific restricted purposes.

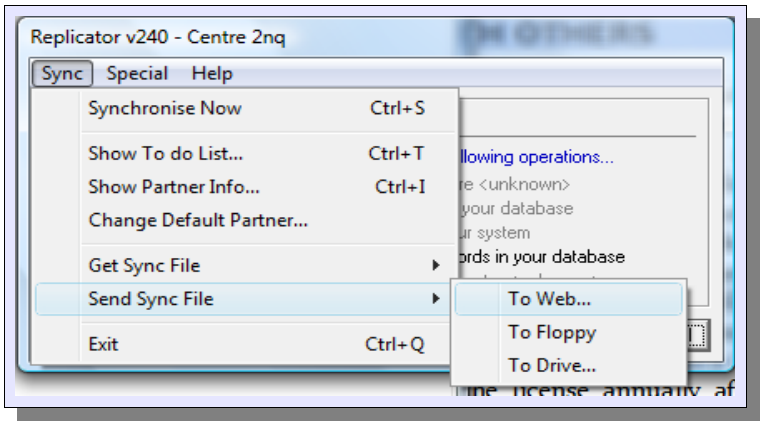
If you have a firewall (and if you use the Internet, you should!) it will have been installed either as part of your *Microsoft Windows* setup, or as part of a third-party security package from a company like Norton or McAfee. Usually, when you first try to make a MapMate transfer via the Web, you will see a screen asking you whether you want to allow the Replicator (referred to as **Repman.exe**) access. You can say 'yes' to this.

On older computers, and with third-party firewalls, you may get rather more obscure problems manifesting themselves. For a number of years I ran a third-party security package on an old version of Windows. Try as I might, I could never persuade the security package that MapMate was a respectable program from one update to the next. The symptoms were that MapMate would go through the motions of uploading to the Web, but would then report that the transfer had been interrupted. If you get this sort of problem, you will need to take it up with your supplier – or else disable the firewall for just as long as it takes to make the transfer.

- **The sync file didn't get generated properly**. You may find that when you come to upload the file that you think you just produced, MapMate can't find it. This is almost always down to a problem with a piece of standard Microsoft software that MapMate uses to compress the file contents, to reduce the time spent transferring the file – imaginatively enough, a program called **compress.exe.**

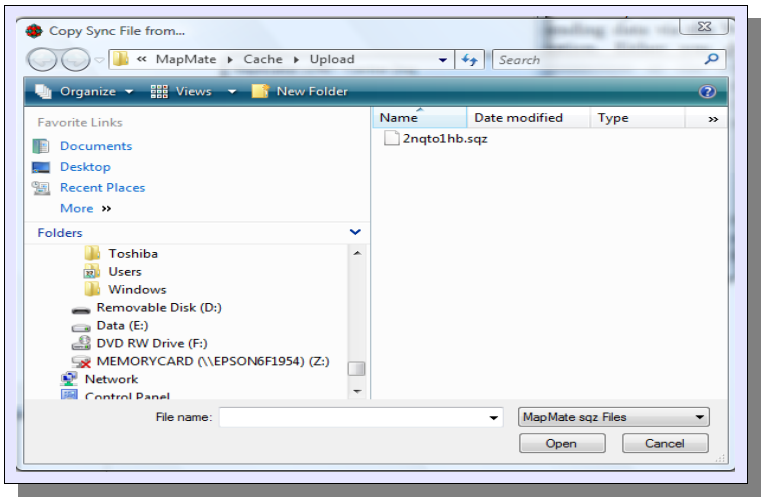
  The remedy is to make sure that you have an up to date version of this program installed on your system. It's available to download for free from the MapMate Web site, and you should follow the current instructions there. I'll talk further about the symptoms of this problem when we cover other ways of sending data below.

Other Ways of Sending Data



There may be circumstances where sending data via the Web is not an option. Either you don't have a connection at the time, or your partner doesn't; or some other factor is making it difficult. In that case you will need to resort to one of the other methods.

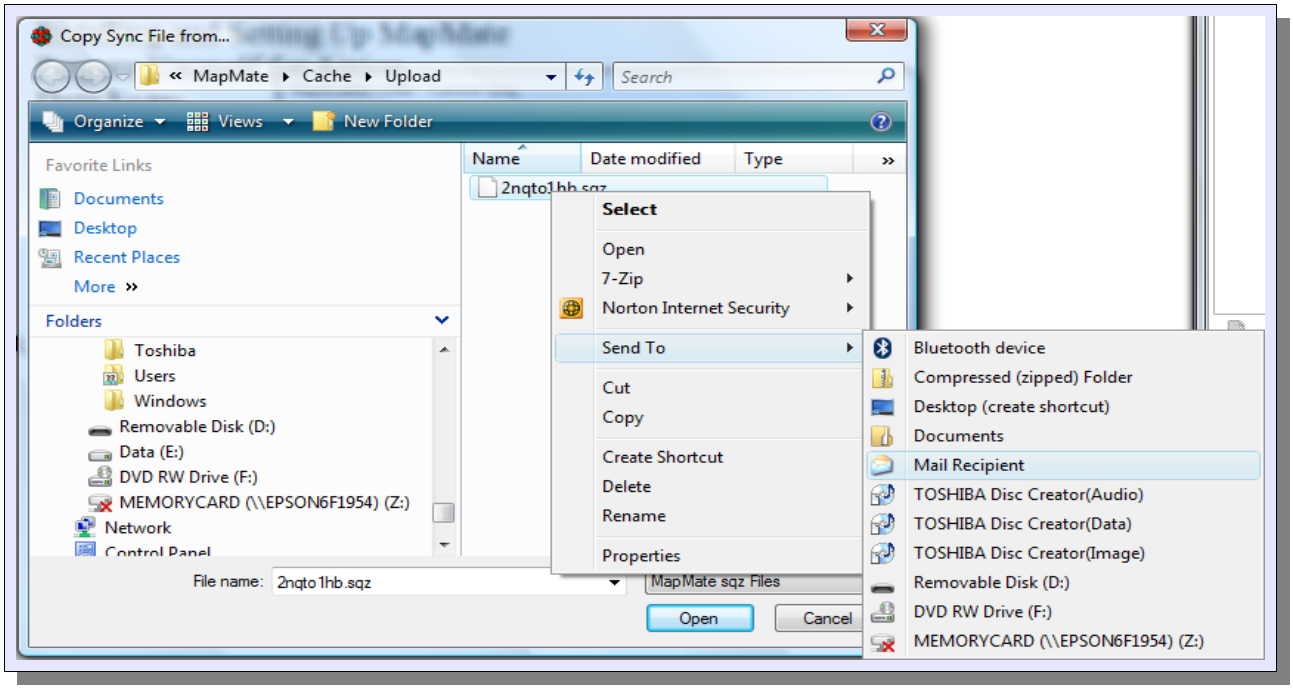
Rather quaintly, MapMate still supports the option **To Floppy**. If your computer still has a floppy disk, you can use this, but you will need to make sure that your partner also has one; also, that you are not sending so much data that it won't fit on a disk.



To Drive... is more useful these days. This will allow you to copy the file to any other storage location accessible from your computer, including CD or DVD drives, removable hard drives, memory sticks, or another location on a local network. You can then make the necessary arrangements with your partner to get hold of the data.

If no file displays on the right-hand side at this point, you may be running an out of date version of the Microsoft data compressor program **compress.exe**. See the earlier section on sending via the Web for how to remedy this.

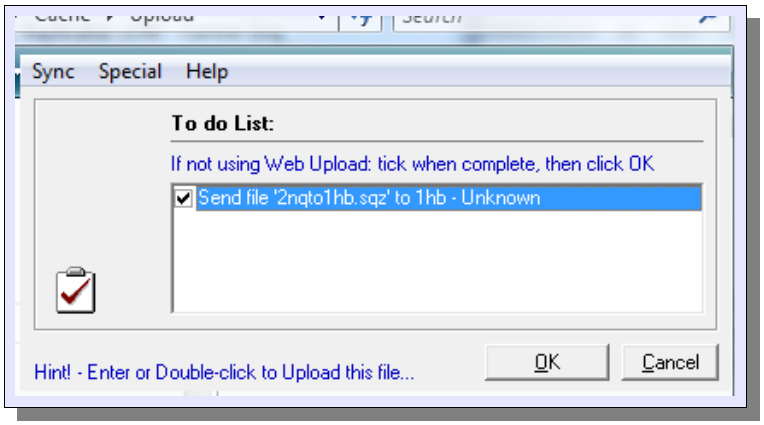
Probably the most useful alternative for sending data is not explicitly listed as an option in the MapMate list; that is, as an email attachment. But you can do it by selecting the **To Drive...** option, **provided** that you have an email client program such as *Microsoft Outlook*, *Outlook Express* or *Mozilla Thunderbird* installed on your computer.



Instead of dragging the file displayed by MapMate to a new location on your computer, right-click on the file name and select **Send To**. Then choose **Mail Recipient**. This should bring up a new message to send in your email program, with the file attached.

Perhaps you handle all your mail through your mail service provider's Web-based mail facilities. In that case, open up your Web browser, log onto your Web email, and use your service provider's facilities for attaching a file to an outgoing message from the MapMate 'Upload' location.

If you are sending a very large data set via email to a partner, remember that some mail service providers impose limits on the size of file that can be transferred, or the total volume of files that can be held on their servers. Happily these limits are getting ever more relaxed. If you do run into trouble, you may be able to overcome the latter constraint by logging on to your service provider's Web site, opening your mail account, and deleting any successfully sent emails that have large attachments.



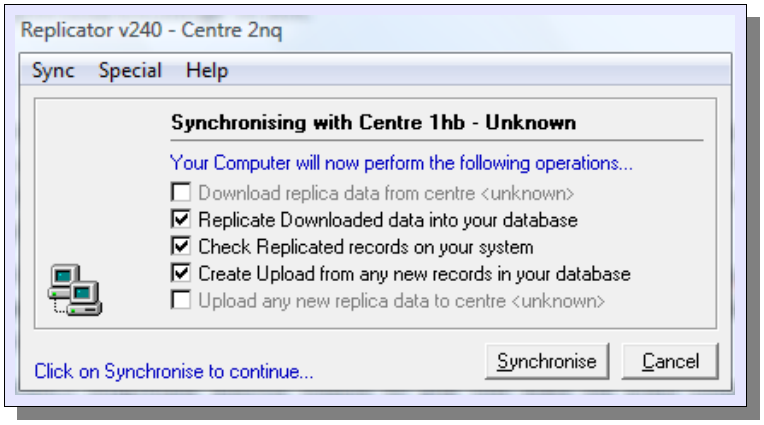
When you send a sync file other than via the MapMate Web service, you will need to go back to the 'To Do' list and update the entry to say you have completed the transfer. Select **Sync/Show To do List...** from the Replicator menu, tick the relevant box or boxes, and click on **OK**. I urge you always to do this immediately; it can be a nuisance if you end up not remembering, when you next want to exchange data with that partner, whether you completed the previous exchange or not.

Receiving Data

Receiving Data From the Web

If you are Internet-connected and a partner chooses to send you exchange data via the MapMate Web server, then you can expect to get an email telling you that it has been sent. Even if you don't see the email, you can always check whether new sync files are waiting for you on the MapMate server, which is what you should do now in any case.

Select **Sync / Get Sync File / From Web...** from the Replicator menu. This allows you to see files that are being held by MapMate ready for your download. Tick all those you want at this time, and set the process going by clicking on **Download**.



MapMate knows where to put the files on your computer; after a successful download all you need to do is return to the Replicator main screen. It will now recognise there are files to deal with, and will display the CUK of the first user in the list. You can then proceed to **Synchronise**.

As with sending records, the process can take a little while, and MapMate will display progress indicators as it processes the file.

Once you have completed a download, you will be asked whether you want to return a sync

file to this partner. If you do send data back to this person and you are both working on updating your own sets, it's a good idea to keep a double-handed exchange going each time one of you sends out something new. However it's not essential to do so immediately, and may sometimes be undesirable – for instance, you may have received a batch of records from a third party that haven't yet been verified.

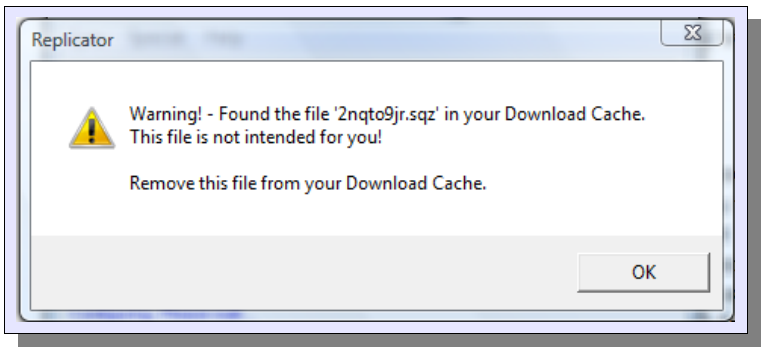
If you choose not to send data back at this time (either because you've now completed an exchange, or you want to hold off for the time being), you can choose whether to send an email saying that you have received the data. I recommend to send either a return file or the email as a matter of good practice and etiquette.

Finally, if you've downloaded files from more than one partner, the Replicator will display the options to replicate data into your system for the next partner in sequence.

Receiving Data By Other Routes

Just as you can send data other than by the MapMate Web service, so you can receive it by the other means already mentioned for sending. But in this case, it is up to you to put the exchange files in the right location for MapMate to find them. If you get a file on exchangeable media you will need to copy it. If you get it as an email attachment you will need to save the attachment. Outgoing files went through an **Upload** folder whose location depends on whether you are running the older or newer regime for MapMate folders and files. Incoming files need to be in the **Download** folder at the same level of the folder hierarchy.

But it's also possible to short-circuit the process of copying the file into the right folder by right-clicking on its name in its original location, and choosing **Open**. This will immediately run the Replicator on that file.



You can verify that the file is really for you, and not sent by mistake, by checking the file name. It should be of the form "<partner CUK>to<your CUK>.sqz"; for instance, as I am user 2nq, I know that a file **1hbto2nq.sqz** is a file for me.

This check doesn't matter too much other than saving you wasted effort, since the Replicator will detect that it's an "invasive non-native" and will not process it.

More on Syncing

What goes on inside a sync file?

This is not strictly something you need to know, but it may help you understand what happens when you need to recover from a tricky synchronisation situation as described in some of the following sections.

A sync file is not simply a copy of the records from your MapMate database. Rather, it's a compressed set of instructions telling the recipient's copy of MapMate what to do with its own database. And it takes a “belt and braces” approach. Suppose that I modify one of the Records on my system. I then sync with you. But my copy of MapMate won't know at that point whether you actually have that Record on your system; you might have killed it off using the 'Remove' facility described in Section 4. Or you might never have loaded an earlier sync. So it includes two sets of instructions: one to modify the data that it thinks might already be there, including all the ancillary data like Sites; and one to add it back in the event that it isn't.

Of course one of these will always fail, but the Replicator will quietly ignore the failure and allow the other instruction to run its course.

The last set of instructions in the sync file updates a special table in MapMate that you have no normal access to. This keeps a record of the sync history with each partner, and adds a new entry for the sending partner which includes details of the last sync sent.

How does MapMate know which updates to include in a sync and which have been sent before? This is down to the generation number; every record in a MapMate system has a generation number attached to it, just as it has a unique identifier. The generation number relates to synchronisation history. When an update or deletion occurs, MapMate puts a marker on its generation number to say that it's a modified or removed record. The Replicator can then check the sync history for a given partner to see whether it's a candidate for the sync or whether it's already gone out since the last update.

Clearly, if my MapMate database has had a more exciting life than yours, the range of generation numbers will be different on my system from yours. But MapMate doesn't try to match our generations; the sending Replicator just recognises that there is something to go out, and the receiving Replicator fits the received updates into the numbering of its own generation history.

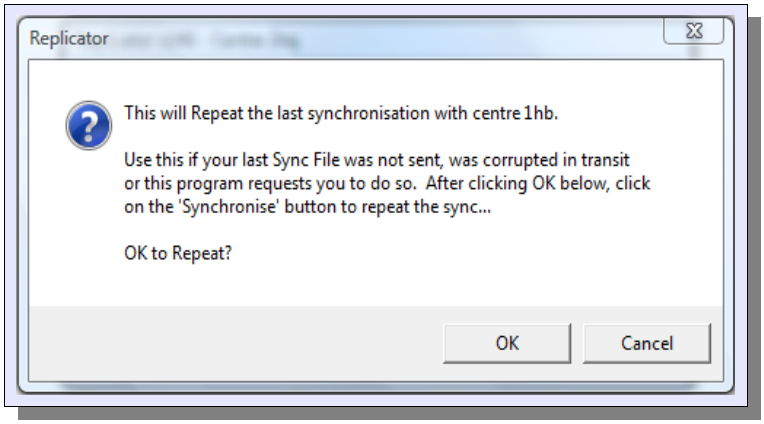
Your Default Partner

If you have a number of partners, the likelihood is that you exchange data with one of them more frequently than with the others. You can use **Sync / Change Default Partner...** to make this the person of first choice when you have no incoming sync files. Use this also on a temporary basis when you want to send to someone else without having had a sync from them.

Resending a Problem Sync

Sometimes a partner may fail to receive a sync file, or may lose track of whether they loaded your last one. In these circumstances, don't simply run another sync file for them.

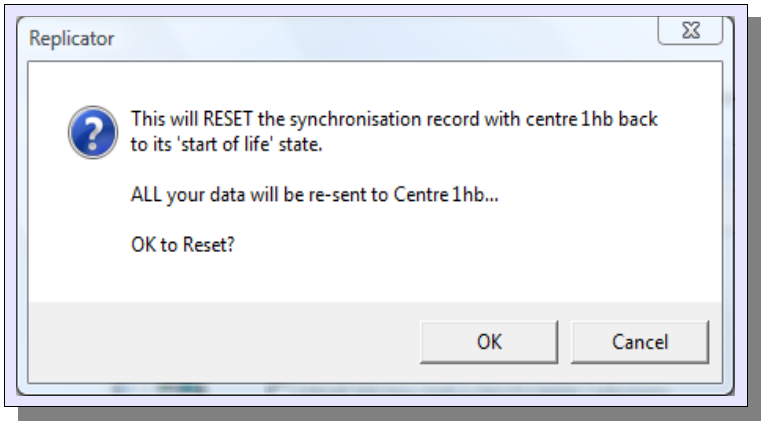
All that will do is transfer any data you collected **since** you generated the previous file, and make it impossible to rewind to the previous set of data.



Instead, make sure they are your current Default Partner, and select **Special / Repeat Last Sync** from the Replicator menu. This will allow you to re-extract all the data you sent last time.

There is no problem with repeating the previous sync more than once. No data will be duplicated as a result.

Resending a Complete Data Set



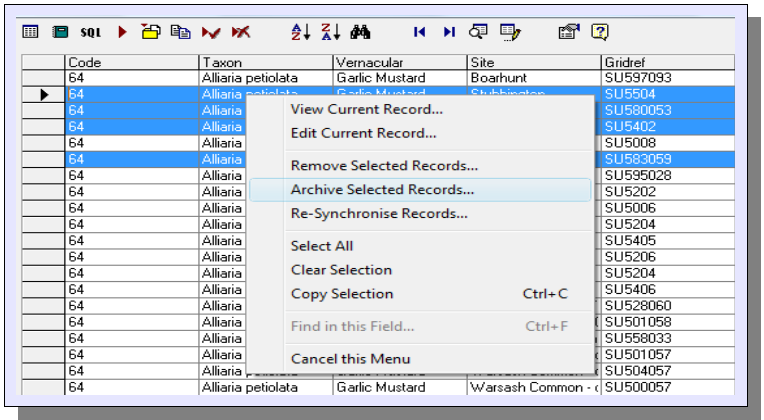
If a partner has a complete disaster with their MapMate system, or possibly gets out of step by more than one data exchange, you will need to resend them a complete update on the data you hold. You can do this by making them your Default Partner and selecting **Special / Reset Sync** Record from the Replicator menu.

Again, apart from perhaps generating a very large sync file that will take a long while to transmit and load, this will have no ill effects, even if they already hold a partial set of your data. No duplication will be introduced.

This is **not** a method of returning their own data to them after a data disaster, even if you do hold a complete set of it from previous exchanges. MapMate will not update any records which have the receiving partner's CUK onto their copy of MapMate. If it did, it would introduce the risk of reverting their data to an earlier state on a normal exchange. To deal with disaster recovery, see Chapter 6 on backup.

Preventing Records from Being Exchanged

If you have certain Records that you don't want to exchange with others, because they are unconfirmed, unverified or have a degree of confidentiality, then you can stop them taking part in any sync data exchanges.



To do this, bring up the Records you want to suppress in the Data Browser grid. Then, using the techniques we

covered in Chapter 4, highlight one or more of them and right-click over the detail of any of them. This is exactly like removing Records, but in this case you will need to select **Archive Selected Records...** 'Archiving' is a rather odd term for this process, as the records remain within your main set of 'current' data.

Should you want to put them back in circulation, you can highlight them again and select **Re-Synchronise Records...** This will then cause them to be sent to each of your partners as you run the next exchange with that partner.

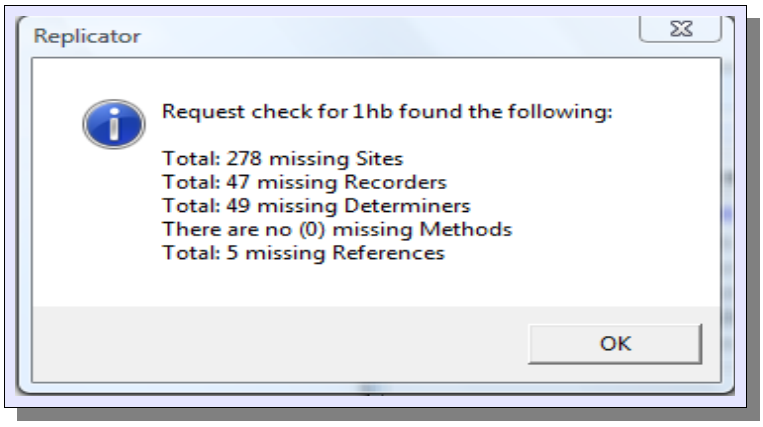
To be frank, this is not a very good way of dealing with confidentiality. It won't stop records being viewable on your computer, or (more seriously) being exported through reports by copying or saving data from the Data Browser. It only affects sync exchanges.

There's nothing within the system's standard reporting tools to indicate which data has been archived, but in Appendix B you will find a query **Browse all archived Records** that will let you find this out.

Missing Data Requests

Occasionally Records lose (or lack) some of their associated data, like Sites or Recorders, on one of the computers that share Record information. When this happens, it may not be readily apparent, because many of the Stock Queries will only display Records for which all the information is complete. If you suspect this may have happened, you can use one or more of the queries provided in Appendix B such as **Browse all Records with Missing Site**, **Browse all Records with Missing Recorder**.

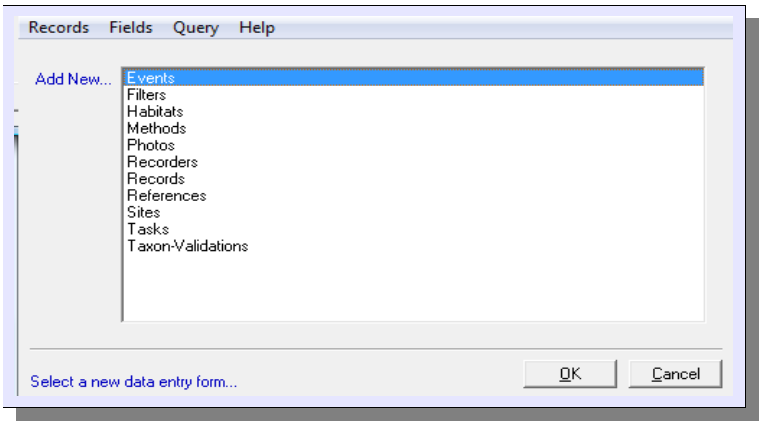
Having determined that missing data is the problem with one or more of a partner's Records, you can ask MapMate to request it on the next sync. Select the person as your Default Partner, then select **Special / Create Missing Data Request** from the Replicator menu.



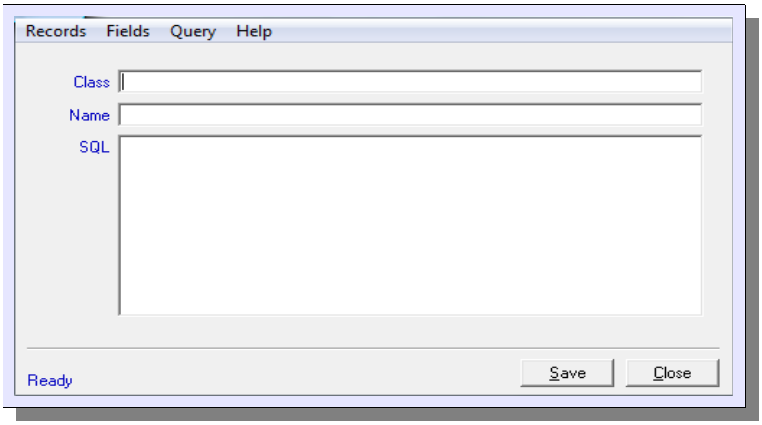
Once you confirm this operation, MapMate will check all that partner's records looking for gaps and display a screen like this. Hopefully you will never see a situation as bad as this, which was specially manufactured for illustrative purposes. You can then **OK** it and proceed with a sync exchange.

Special Filters

When we were setting up filters for exchange partners, you may have wondered whether you can provide more specialised, customised filters. The answer is 'yes', but you will have to apply some of the more specialist skills you may have acquired in Chapter 4 for custom reporting. Let's suppose you have a partner who is interested in having only records for true Sedge (*Carex*) species from you.



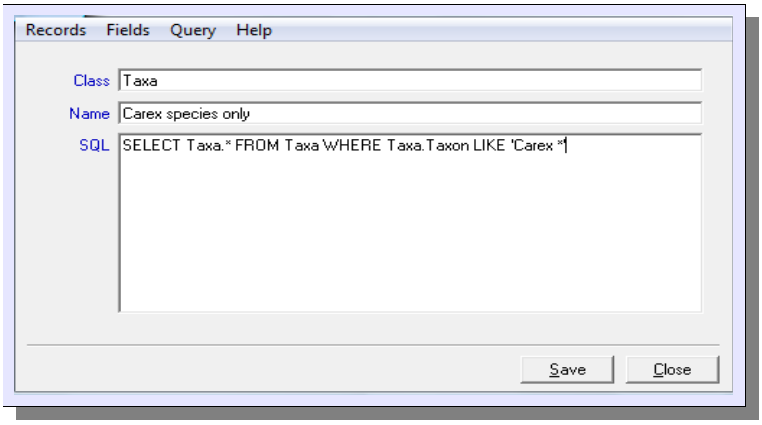
Open the Data Entry form from the MapMate main menu, and from **Records** select **Add New...** You should see **Filters** in this list. If is not, you will first need to download and install patch 278 from the MapMate Web site. Chapter 10 has advice on downloading patches, if you haven't done this before.



Select **Filters**, which will bring up this form. There are three possible entries for **Class**, which are **Records**, **Sites** and **Taxa**. One of these has to be typed into the editing box. In this case, it is **Taxa**.

The **Name** entry is whatever you want displayed as an aide-memoire for your filter. If you're not certain whether a name is already in use, click on the **Class** caption and select **Browse Class...** This will bring up a Data Browser for all Filters, where you can search for names. In this case, enter **Carex species only** as the **Name**.

The SQL box is the tricky one, and as for report queries, you must type perfectly correct computer language statements into this box, with no spelling or punctuation errors. For a **Taxa** filter the form of the query will be:



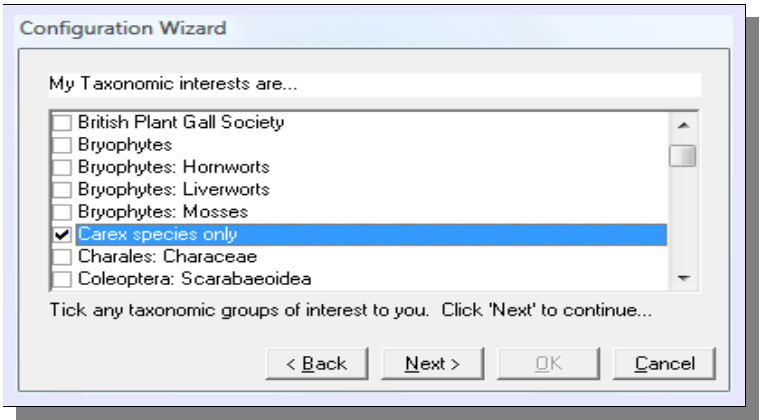
SELECT Taxa.* FROM Taxa WHERE <criteria for selection>

In the current example (note the space after **Carex**),

SELECT Taxa.* FROM Taxa WHERE Taxa.Taxon LIKE 'Carex *'

If you followed the examples in Chapter 4, note that there is no semi-colon at the end of the query in this context.

Save this completed example by clicking on the **Save** button. When the form comes back empty, you can click on **Close**.



That's created the Filter, but you still have to make it available for use. To do this, return to the MapMate main screen and select **View / My**

Configuration... Go to the second screen of the Configuration Wizard, search for **Carex species only**, and tick its box.

You can then continue through the rest of the configuration process keeping the same settings as when you set up MapMate.

You will now find that you can use **Carex species only**, not only for filtering data exchanges with partners, but also for setting Defaults when carrying out data entry or reporting.

As an exercise, you might like to try setting a **Sites** filter which selects only Sites in the hectad SU50. The general form for a **Sites** SQL query is

SELECT Sites.* FROM Sites WHERE <criteria for selection>

We won't cover **Records** filters here, which usually require more complex queries.

Text Import

When Is Text Import Appropriate?

The exchange mechanism we've discussed so far is fine for sharing data with other MapMate users who have a commitment to maintaining their own records. Of course not everyone who passes records on to you will have MapMate, even when they have a computer. Some of those who do have MapMate may simply want to use it as a vehicle for letting you have the initial record, and prefer to leave any further management of the data in your hands.

For the two latter groups of people, the best way to get their records in electronic form is via text import. This is a “one off” transfer of data that leaves you with ownership of the imported records.

Sometimes a person who has been exchanging records with you via MapMate syncs wishes to pass the responsibility for future maintenance of their data over to you – perhaps because they are moving out of your area, or for health reasons. In this case, the best course is to take over that person's editing rights. This is something that you need to request from MapMate Ltd. technical support; they will then send you a patch that enables you to edit any data that has the other person's CUK attached.

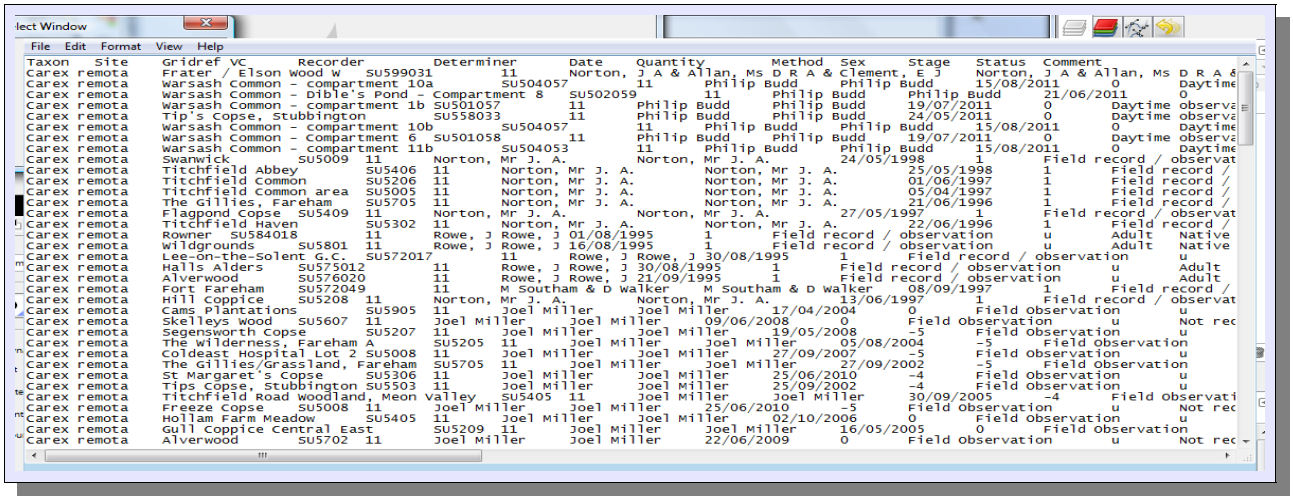
A more cumbersome approach is that they generate a text export of all their records, save it to a file, delete all their records, perform a last sync with you to remove their original records from their system, then pass you the text file to import. Apart from being more cumbersome and carrying a slight element of risk, this has the additional disadvantage that some detail may be lost from their data, as explained below. If you leave out the deletion stage in this process, you will end up with duplicates of all their records, one set owned by them and one set by you.

The Text Import File Format

Data brought in by text import must be in a strict format. The file itself is “plain text” with a very limited set of adornments; so a normal word-processing document won't do. Each line in the file represents one Record, and there can be no line breaks anywhere in the middle of a Record. Each item in the Record is separated from its neighbours by what is known as a “tab” (tabulator) character, which (for those of us old enough to remember typewriters) works like the tabulator on a typewriter, leaving sufficient white space to move the printing on to a fixed position to the right – though for current purposes it only matters that it separates items of data. For that reason, no item of data can have a tab character in the middle; and there should only ever be a single tab between items, otherwise the MapMate importer will think there is an extra, blank data item in the Record. Also, there should be no blank lines anywhere in the file, or at the end of the file.

If you're curious, this is what part of an import file looks like in a plain text editor (*Microsoft Notepad*). If you really wanted to, you could enter Records using an editor like this, but it would clearly be very awkward. Because data items are of different lengths in

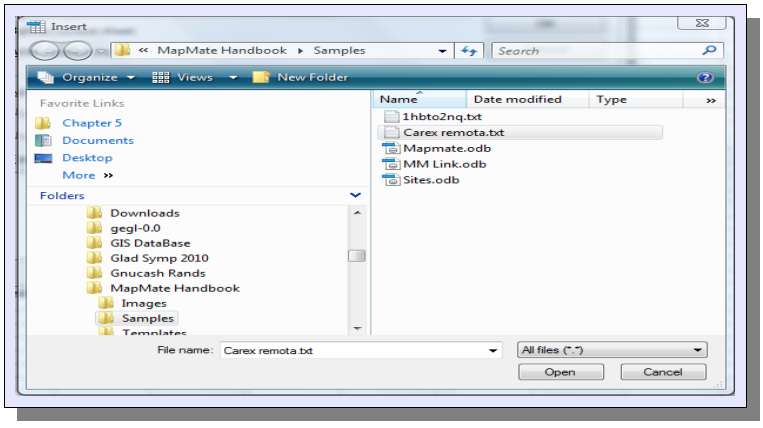
different Records, and there can only be a single tab between each item, nothing lines up neatly.



Notice, however, that there is a first row that contains all the **names** of the data items, followed by rows representing the Records with their data values.

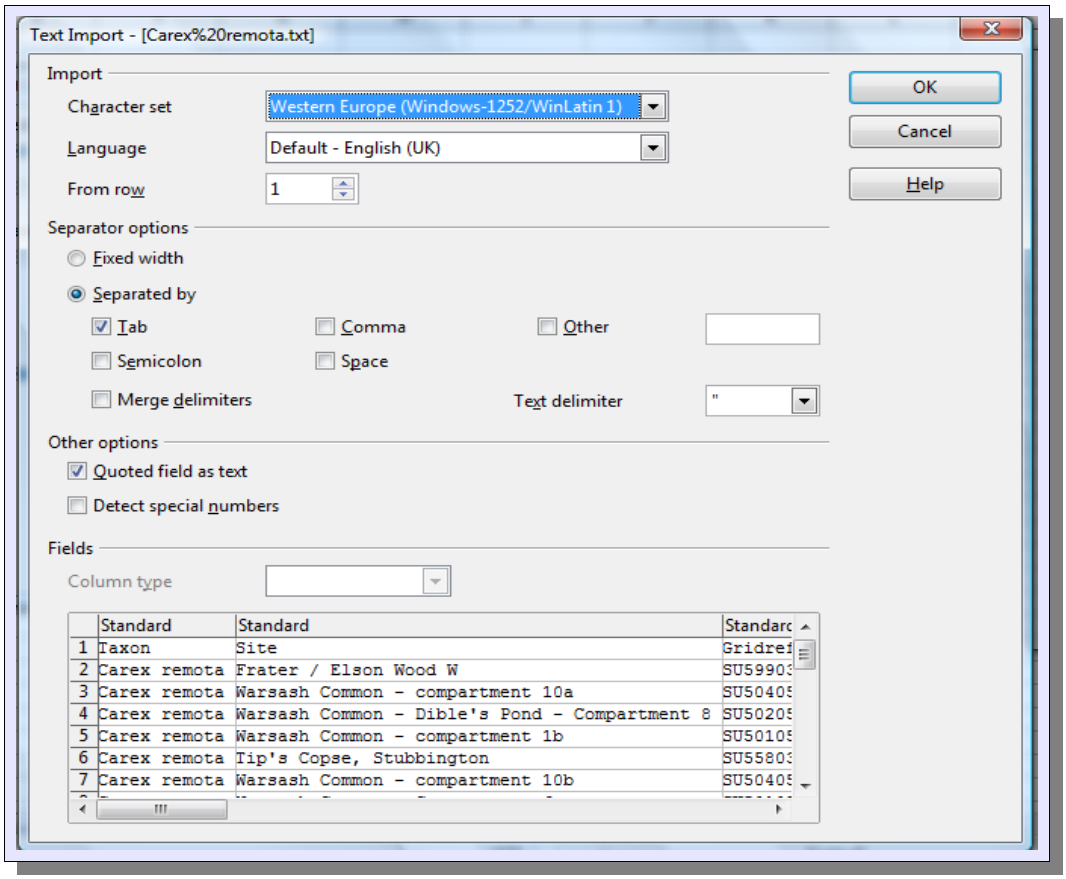
Fortunately, this type of file (often called, unsurprisingly, "tab-delimited text") is quite a commonly supported format. Many programs that are easier to use than a plain text editor will allow you to read and write these files. The most convenient type of program for limited manual data entry is a spreadsheet, and I'll base my examples on this from now on. But if one of your partners is running another database program to keep their biological records, they should be able to find a way to export the data in the required form.

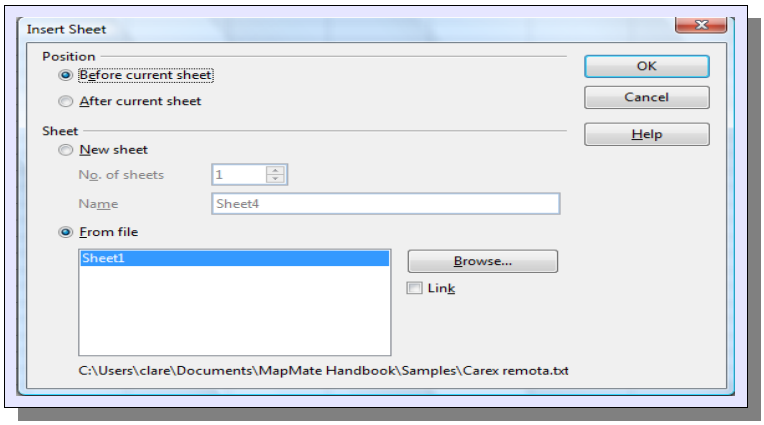
When you get such a file from someone else, you may well need to check that it's in an appropriate form, so we'll start by looking at putting one into an *OpenOffice Calc* spreadsheet. *Excel* users will have a slightly different procedure, but the principles are the same.



With a blank spreadsheet open, select **Insert / Sheet from File...** from the *Calc* main menu bar. Locate and open the file from your computer.

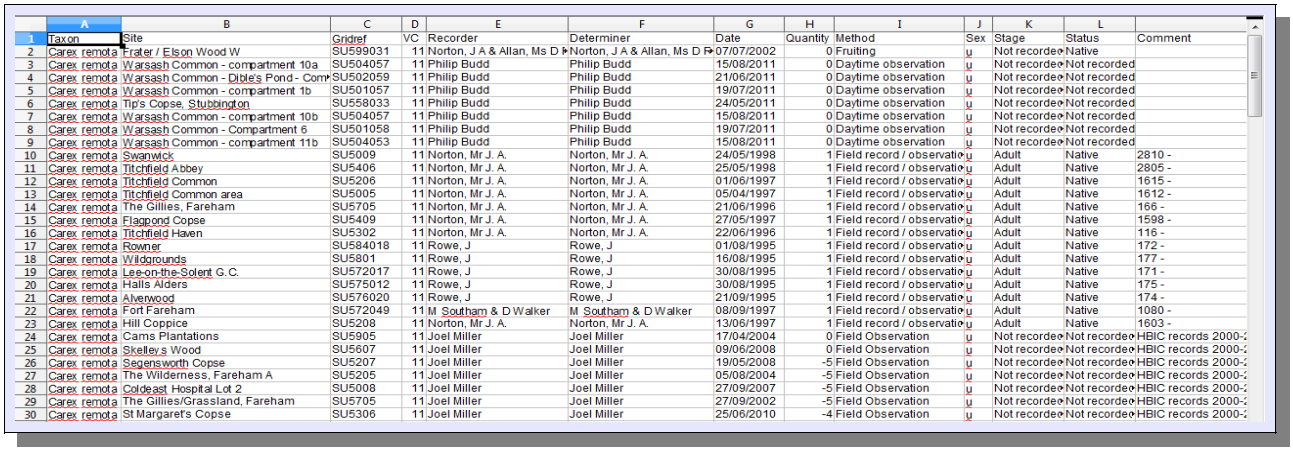
The main thing on the following screen is to make sure that the spreadsheet program has correctly deduced the format of the file. If it hasn't, adjust the settings. Make sure that **Separated by** is selected rather than **Fixed width**, and that the only box ticked under that heading is **Tab**. The sample display at the bottom of the form will give you a clue; check that the columns line up properly with the first few titles as expected. Then you can click **OK**.





Calc will then show you a further screen, as it inserts your data in a separate sheet in the spreadsheet workbook. It doesn't really matter where it goes, so you can accept the defaults at this point and just click **OK**. *Excel* works differently, by inserting data into the sheet you're currently looking at.

At this point you should see a display like that below. I've narrowed some of the columns so that all the columns show up, at least in part. The first row should have column names spelt exactly as shown, and as listed in the account below.



	A	B	C	D	E	F	G	H	I	J	K	L	
1	Taxon	Site	Gridref	VC	Recorder	Determiner	Date	Quantity	Method	Sex	Stage	Status	Comment
2	Carex remota	Frater / Elson Wood W	SU599031	11	Norton, J A & Allan, Ms D F	Norton, J A & Allan, Ms D F	07/07/2002	0	Fruiting	u	Not recorde	Native	
3	Carex remota	Warsash Common - compartment 10a	SU504057	11	Philip Budd	Philip Budd	15/08/2011	0	Daytime observation	u	Not recorde	Not recorded	
4	Carex remota	Warsash Common - Dible's Pond - Com	SU502059	11	Philip Budd	Philip Budd	21/06/2011	0	Daytime observation	u	Not recorde	Not recorded	
5	Carex remota	Warsash Common - compartment 1b	SU501057	11	Philip Budd	Philip Budd	19/07/2011	0	Daytime observation	u	Not recorde	Not recorded	
6	Carex remota	Tip's Copse, Stubbington	SU558033	11	Philip Budd	Philip Budd	24/05/2011	0	Daytime observation	u	Not recorde	Not recorded	
7	Carex remota	Warsash Common - compartment 10b	SU504057	11	Philip Budd	Philip Budd	15/08/2011	0	Daytime observation	u	Not recorde	Not recorded	
8	Carex remota	Warsash Common - Compartment 6	SU501058	11	Philip Budd	Philip Budd	19/07/2011	0	Daytime observation	u	Not recorde	Not recorded	
9	Carex remota	Warsash Common - compartment 11b	SU504053	11	Philip Budd	Philip Budd	15/08/2011	0	Daytime observation	u	Not recorde	Not recorded	
10	Carex remota	Swanwick	SU5009	11	Norton, Mr J. A.	Norton, Mr J. A.	24/05/1998	1	Field record / observatio	u	Adult	Native	2810 -
11	Carex remota	Titchfield Abbey	SU5406	11	Norton, Mr J. A.	Norton, Mr J. A.	25/05/1998	1	Field record / observatio	u	Adult	Native	2805 -
12	Carex remota	Titchfield Common	SU5206	11	Norton, Mr J. A.	Norton, Mr J. A.	01/06/1997	1	Field record / observatio	u	Adult	Native	1615 -
13	Carex remota	Titchfield Common area	SU5005	11	Norton, Mr J. A.	Norton, Mr J. A.	05/04/1997	1	Field record / observatio	u	Adult	Native	1612 -
14	Carex remota	The Gillies, Fareham	SU5705	11	Norton, Mr J. A.	Norton, Mr J. A.	21/06/1996	1	Field record / observatio	u	Adult	Native	166 -
15	Carex remota	Flagpond Copse	SU5409	11	Norton, Mr J. A.	Norton, Mr J. A.	27/05/1997	1	Field record / observatio	u	Adult	Native	1598 -
16	Carex remota	Titchfield Haven	SU5302	11	Norton, Mr J. A.	Norton, Mr J. A.	22/06/1996	1	Field record / observatio	u	Adult	Native	116 -
17	Carex remota	Rowner	SU584018	11	Rowe, J	Rowe, J	01/08/1995	1	Field record / observatio	u	Adult	Native	172 -
18	Carex remota	Wildgrounds	SU5801	11	Rowe, J	Rowe, J	16/08/1995	1	Field record / observatio	u	Adult	Native	177 -
19	Carex remota	Lee-on-the-Solent G.C.	SU572017	11	Rowe, J	Rowe, J	30/08/1995	1	Field record / observatio	u	Adult	Native	171 -
20	Carex remota	Halls Alders	SU575012	11	Rowe, J	Rowe, J	30/08/1995	1	Field record / observatio	u	Adult	Native	175 -
21	Carex remota	Alverwood	SU576020	11	Rowe, J	Rowe, J	21/09/1995	1	Field record / observatio	u	Adult	Native	174 -
22	Carex remota	Fort Fareham	SU572049	11	M Southam & D Walker	M Southam & D Walker	08/09/1997	1	Field record / observatio	u	Adult	Native	1080 -
23	Carex remota	Hill Coppice	SU5208	11	Norton, Mr J. A.	Norton, Mr J. A.	13/06/1997	1	Field record / observatio	u	Adult	Native	1603 -
24	Carex remota	Cams Plantations	SU5905	11	Joel Miller	Joel Miller	17/04/2004	0	Field Observation	u	Not recorde	Not recorde	HBIC records 2000-2
25	Carex remota	Skelley's Wood	SU5607	11	Joel Miller	Joel Miller	09/06/2008	0	Field Observation	u	Not recorde	Not recorde	HBIC records 2000-2
26	Carex remota	Segensworth Copse	SU5207	11	Joel Miller	Joel Miller	19/05/2008	-5	Field Observation	u	Not recorde	Not recorde	HBIC records 2000-2
27	Carex remota	The Wilderness, Fareham A	SU5205	11	Joel Miller	Joel Miller	05/08/2004	-5	Field Observation	u	Not recorde	Not recorde	HBIC records 2000-2
28	Carex remota	Coldeast Hospital Lot 2	SU5008	11	Joel Miller	Joel Miller	27/09/2007	-5	Field Observation	u	Not recorde	Not recorde	HBIC records 2000-2
29	Carex remota	The Gillies/Grassland, Fareham	SU5705	11	Joel Miller	Joel Miller	27/09/2002	-5	Field Observation	u	Not recorde	Not recorde	HBIC records 2000-2
30	Carex remota	St Margaret's Copse	SU5306	11	Joel Miller	Joel Miller	25/06/2010	-4	Field Observation	u	Not recorde	Not recorde	HBIC records 2000-2

Column Contents

Taxon This needs to be the name of a Taxon that is recognised by MapMate and spelt **exactly** the same. This includes abbreviations like 'agg.' and 'var.'. If you are not sure, you can try looking it up through the MapMate Data Entry form or run the **Browse Taxa** report. Alternatively, you can wait until you load the file into MapMate and let the Importer tell you about any unrecognised names – but in that case, you will need to go round the editing loop again; you can't amend the file within MapMate.

Site This one is pretty straightforward, but there is a limit of 64 characters on a Site name. If there's extra descriptive detail in here that takes it beyond that limit, you can consider removing it from here and adding it to the Comment entry. See also the strictures about undesirable formatting characters in the notes on the Comment.

Gridref Check this one carefully. 10-figure (1-metre) grid references do not work on some earlier versions of Mapmate; they accept them, but then mangle the reference. No spaces are allowed anywhere in the grid reference. Tetrad references in 'DINTY' format (e.g. **SJ47G**) are acceptable and should always be used for records at tetrad resolution; importing this reference as **SJ4272** will give a spurious 1-km resolution. While you're looking, you may as well check for 100-km square letters that are off-beam, and reversed Eastings and Northings.

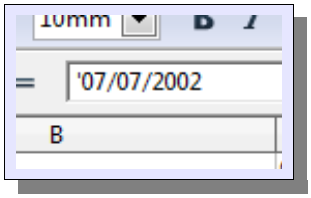
VC Should be the simple vice-county number. Irish vice-counties should be in the form '200 + VC number' discussed in Chapter 4, rather than 'H + VC number'.

Recorder The limit on a Recorder name is again 64 characters. You may have your own standards for whether and how recorder names are laid out consistently. With multiple Recorders, all names must go in the single entry.

Determiner Determiner can be left blank, and in that case MapMate will interpret the entry as 'same as Recorder'. Otherwise, the same considerations apply as to Recorder.

Date This is another one to check carefully. The date **must** be in the format 'dd/mm/yyyy'; for example **13/06/2007**. Only forward slashes can be used as separators. The year must be in four-digit format. MapMate will accept arbitrary date ranges, but they must be in the form of two dates specified in this way, separated by a hyphen and with no spaces around the hyphen; e.g. **01/06/2007-14/06/2007**. Alternatively, you can specify a month by setting the calendar day to '00' (make sure it's two digits); for example, 'August

2011' would be **00/08/2011**. You can go further and specify a year only by making the month '00' as well. So a date of 1974 would be **00/00/1974**.



If you followed the steps I suggested with *Calc*, you'll find that the Date items in your spreadsheet have been imported as literal text, rather than a value to be interpreted as a date. You can tell this by putting your cursor into one of the Date cells on the spreadsheet; the value for editing is shown in the toolbar with a tick mark to its left, indicating that this is text pure and simple. If you use a different program or load the data differently, there's a possibility that these items will have been interpreted as dates, and given a different internal representation in the program. In that case, compound dates, and dates with zeros in the day or month part, will return weird results. You'll find that your spreadsheet program, when letting you specify the file to import, will also let you override the interpretation of any column. If you get this undesirable effect, I suggest you re-import the data forcing this column to be 'Text'.

If you change the contents of any of the date cells after importing the file, you may also find that the cell changes from being a text field to a date field – perhaps changing the date format in the process. To prevent this happening, you can highlight the whole Date column on your spreadsheet, select **Format / Cells...**, then change the current format to **Text**. On *Calc* this actually makes the tick mark in the cell editing box disappear, but don't worry; *Calc* still recognises it as text.

Quantity This can be blank or zero, in which case the Importer will interpret it as 'Present'. 'Not Present' or 'Not Found' must be represented as -7, and DAFOR scores should be given codings in the range -1 to -6, -21 to -23 as described in Chapter 4. Positive numbers represent actual numbers, and the Comment will need to say whether these are exact counts or estimates if that is significant. You can't use any of the tricks described for data entry under Chapter 3 to add extra stock commentary to the Comment.

Method This can be left blank, in which case it will be interpreted as 'Unknown'. If the text in this item exactly matches a Method name already in MapMate, it will link to that. Otherwise a new Method will be created. This is generally a nuisance, as a standard set of Methods is usually what you want, and you don't want near-duplicates or totally inappropriate categories of information put into here. (You may have noticed that one such has crept into the very first line of our imported file, illustrated above.) The Importer also allows you to check for a range of novelties introduced by an import file, so you can choose either to do it by eye here or to wait for it to be picked up later.

Sex Can be left blank or assigned 'u' for 'Unrecorded', or assigned 'm' or 'f.

Stage This should be one of the recognised Stage terms in MapMate, or left blank for 'Not recorded'.

Status Similarly, this should be one of the recognised Status terms in MapMate, or left blank for 'Not recorded'.

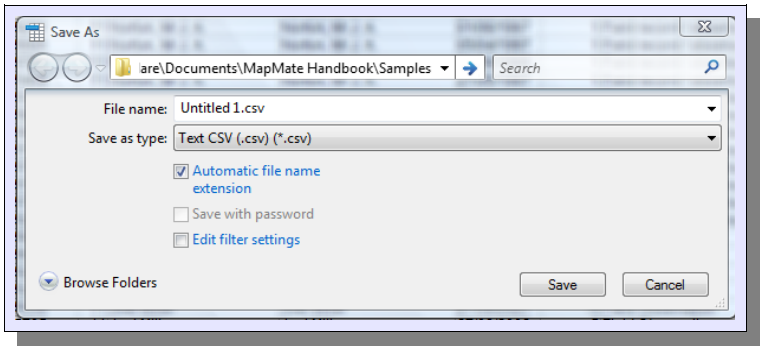
Comment The Comment may be blank, or can be reasonably free text of considerable length, but it's worth checking that the originator has not let any special formatting characters slip in. A line break in the middle of a Comment will result in the

next row in your spreadsheet starting unexpectedly with text that isn't a Taxon name, and everything else in the row out of kilter. A tab character in the middle of a Comment will probably result in the Comment being split across two columns, the normal Comment column and the unlabelled one to its right. Other odd characters may show up as odd symbols not looking like letters, digits or normal punctuation marks, and should be treated with suspicion.

I advocate that if you see problems of this sort, you get the originator to fix them at source and re-send you the file. It helps to ensure that they are conscious that a problem is caused, and that reduces the risk of it happening again. If this isn't practical, you may want to fix the problem within the spreadsheet.

Saving an Edited Import File

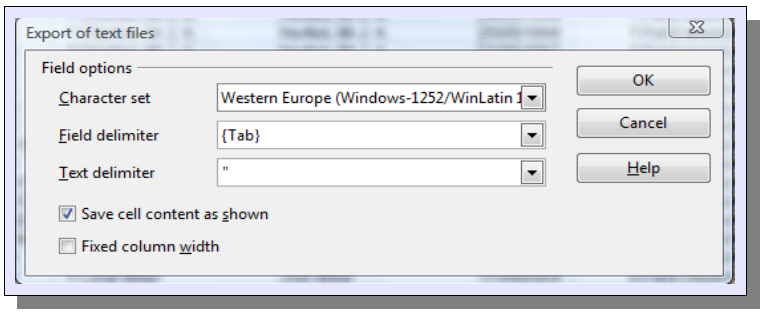
If all was well with the original data imported into the spreadsheet, you can close the spreadsheet down without saving it (but of course, hang on to the original text import file!). If you made changes, you will need to save the spreadsheet as a new tab-delimited file. I recommend that you save it with a different name from the one you imported, as that means you can go back to the original if you did something silly during your edits.



In OpenOffice Calc, that means selecting **File / Save As...** from the main menu bar. Then, from the drop-down **Save as type** list, select **Text CSV**. If you know about these things, that may come as a bit of a surprise, as CSV format files are actually a different format of text file. But don't worry.

Select the location where you want to save it on your computer, and change the file name to something memorable, adding the file extension **.txt** to the name. Uncheck the box labelled **Automatic file name extension**. Then click on **Save**.

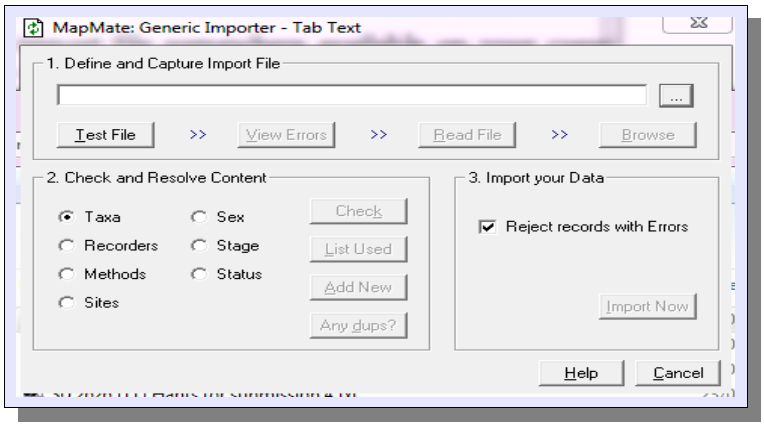
You will probably get a warning saying that this file format doesn't support all the wonderful features your spreadsheet program has to offer. You can gaily accept that (**Keep Current Format** in *Calc*).



Make sure that on the next screen, the **Field delimiter** is **{Tab}** and the **Text delimiter** box is empty; make sure to delete the double quote (**"**) that is offered by default. Then click on **OK**. You may get another warning that only the current sheet is saved; that's OK too.

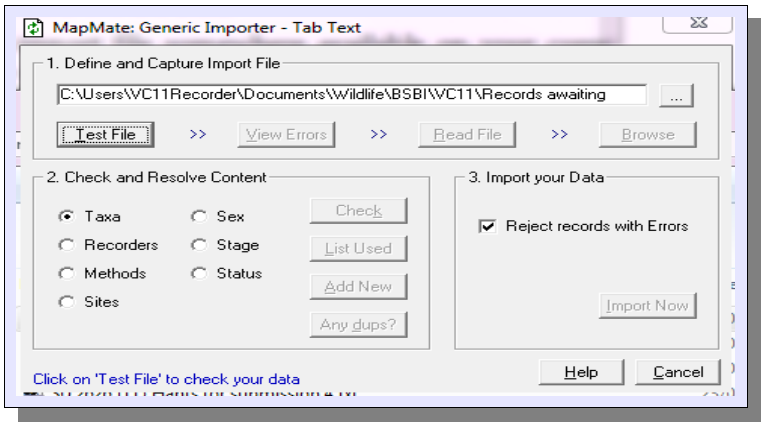
You can now move on to the actual import into MapMate.

Running the Import Process



Once you have the text import file somewhere available on your computer, select **File / Import** from the MapMate main menu. Then select **Data from Tab Text Files...**

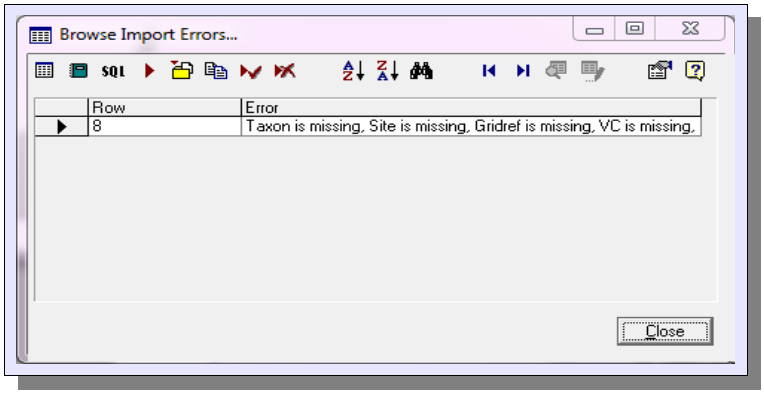
First click on the button with the ellipsis (...) to the right of the long data entry field at the top. This will allow you to browse your computer and select a file in the usual Windows way. By default, MapMate will look for files with the extension '.txt' on their name, but if your file ended up with another extension, you can widen the search to look for that. What matters is that the file format and content meet MapMate's specification.



Having selected your file, you will then need to click on the button **Test File** to move the process on. Hopefully MapMate will tell you that all is well, but you may see one of two messages: either **File has Fatal Errors(s). Header Row is Not to Specification – See Help** or **File has Errors. Click on the 'View Errors' Button for details.**

In either case, you will need to go back to the import file and correct it. Simple errors in the header or in one or two lines of the file can be tackled with a plain text editor like Windows *Notepad*, if you feel confident in using one. More extensive or systematic problems are best tackled by reloading the import file back into a spreadsheet program where it is easier to see the problem on a regular layout and make bulk changes.

If you get the former of these messages, then one or more captions are missing from the first line of your file, or they are mis-spelt, or they are not properly separated by tab characters. (The last of these problems may not be easy to spot in *Notepad*, as tabs and multiple spaces will look exactly the same.)



If you see the latter message, you should follow the instructions, which will reveal where the problems are. Most problems are fairly self-explanatory, but the one shown here may cause a bit of head-scratching. It is in fact caused by one or more blank lines within or (more likely) at the bottom of your import file.

If your import file is generated from Microsoft *Excel*, this is something that can creep in rather easily. OpenOffice *Calc* doesn't seem to have the same problem, as it always omits trailing blank rows from text export files.

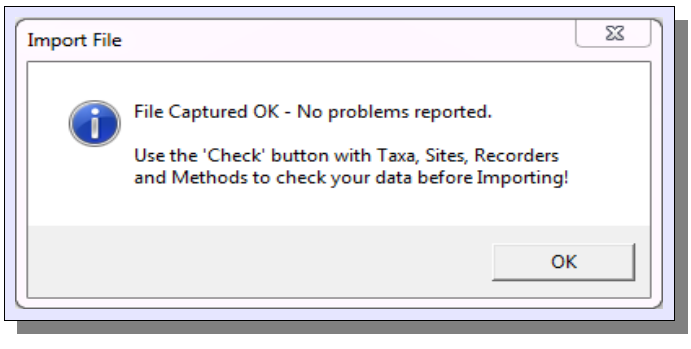
If you use *Notepad* to re-edit your file, the blank line may not be very easy to spot, as there is no line numbering and the final blank line doesn't actually get displayed. However, if you can put your cursor in a blank line below the last text, you should use the **Backspace** key to delete this final line break.

```
1  Taxon Site  Gridref VC  Recorder  Determiner  Date  Quantity  Method  Sex Stage Status  Comment
2  Lamium maculatum  Highbridge Farm (main): Highbridge  SU467211  11  David S Hubble  David S Hubble  24/03/2012  -5          1 patcl
3  Luzula campestris Houghton: Bank by churchyard  SU341326  11  Martin W Rand Martin W Rand 22/03/2012  0
4  Erophila verna  Flexford: Caravan site  SU421221  11  Martin W Rand Martin W Rand 23/03/2012  -22
5  Picea abies Flexford: Caravan site  SU421221  11  Martin W Rand Martin W Rand 23/03/2012  0          Planted
6  Picea sitchensis  Flexford: Caravan site  SU422220  11  Martin W Rand Martin W Rand 23/03/2012  -5          Planted
7  Viscum album  Oak Close, Upham: By Upham Street SU536204  11  Martin W Rand Martin W Rand 25/03/2012  0
8  Viscum album  Winchester Road, Bishops Waltham: SU5318  11  Martin W Rand Martin W Rand 25/03/2012  -5
9
10
```

Here I'm illustrating the problem using a rather more sophisticated editor, *Bluefish* (another free, open-source program that you can download from the Internet). My cursor is at the start of a line ready to create a "line 10", but between this and the text is an empty "line 9". The MapMate importer flagged this as "row 8", because it's not counting the header. I can now delete this blank line.

If you choose to fix the problem in *Excel*, then the way to get rid of blank lines at the end of the document is to place your cursor in the left-hand cell of the first blank line at the bottom. Then hold down **Ctrl** and **Shift** simultaneously with pressing the **End** key. If (and only if!) this causes the cursor to move down into a position in another blank row, or rightwards across the current row, you should then select **Edit** / **Delete** from *Excel*'s main menu. The likelihood is that this will pop up a window asking you what and how you want to delete. The option to choose is **Rows**, then click on **OK**. Once that is done, you can save the sheet as a tab-delimited text file again.

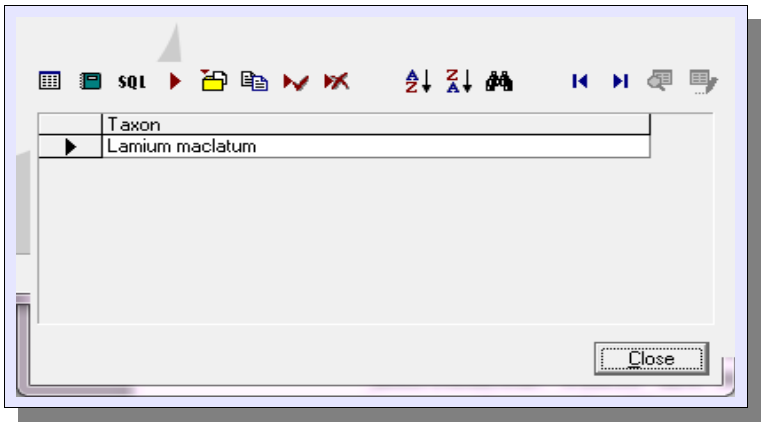
Whatever program you use to make corrections, you should both save and close the file once you've done it. MapMate may object to re-loading a file that is still open in another program, even if changes have been saved. Curiously, this doesn't seem to apply to all programs.



Once you have eliminated these sources of problems, you should go back to **Test File** to force a reload. You can then proceed to **Read File** and, all being well, you will see this box. If you see a warning that the file may have changed and is not being re-read, go through **Test File** and **Read File** once more and the problem will disappear. If you wish, you can then use the **Browse** button to see what you're getting. In practice, this is not a very useful facility as it doesn't let you correct any inaccuracies in the data you're receiving, or remove any Records you don't want. If you have a large file, it's probably best to move on to the next stage.

You now need to test the various data items to see whether any unwanted novelty will be introduced by the import. How many of these items you test depends on how stringent you

are in allowing things such as new Recorder names or new recording Methods to be introduced by this route. There are some items that you must check if you don't want to end up with rejected records. They are **Taxa**, **Sex**, **Stage** and **Status**. Other items such as **Site** or **Recorder** can be left, if you choose, and the import process will automatically define new entries for these if they don't exist.

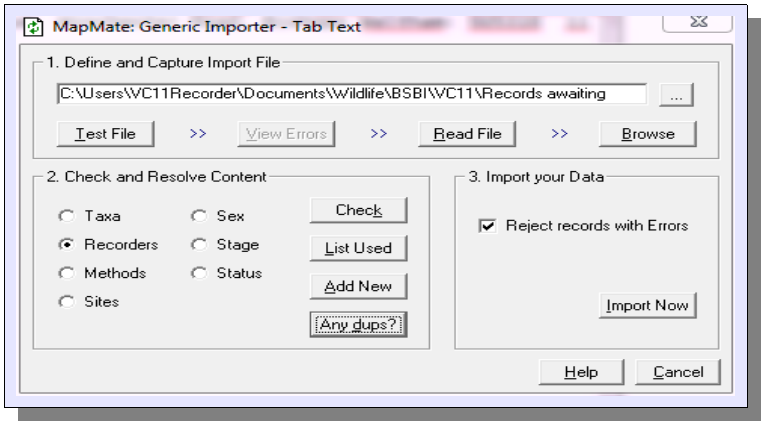


Taking **Taxa** as an illustration, click on the **Taxa** radio button and then click on **Check**. Hopefully this will report that all Taxa are already defined in MapMate. If it doesn't, you will see a grid display of the records at fault. Here the error is fairly clearly a spelling mistake. You will need to edit the file, and again test and read it. Similar problems will arise if the supplier of the data has used different taxonomy or nomenclature from that recognised by MapMate, and you may have to dig a little deeper.

If you want to see exactly what items of data are being brought in with this import, check the radio button for the category of data and click on **List Used**. Again, you can't make changes from this list.

However you can pre-empt the automatic definition of a new entry from the import by using the **Add New** button, again with the appropriate radio button checked. The import process brings in a limited amount of information for each category of data, and MapMate will provide defaults for the rest. If you want to set up anything other than this, use this facility, making sure that all the relevant detail (for instance, Site Name, Gridref and VC for a Site) match exactly those shown in the import file.

If there is an entry in the MapMate database matching the detail supplied in the import file, there's always the possibility that more than one entry matches. In that case, MapMate will pick up one of the possible matches in an arbitrary fashion. To find out whether this is a possibility for detail in any of the Records (for instance, Recorders with the same name more than once on MapMate), select the category on the radio buttons and click on **Any dups?** This will display any instances of duplicates already on MapMate that could be targeted by the import.

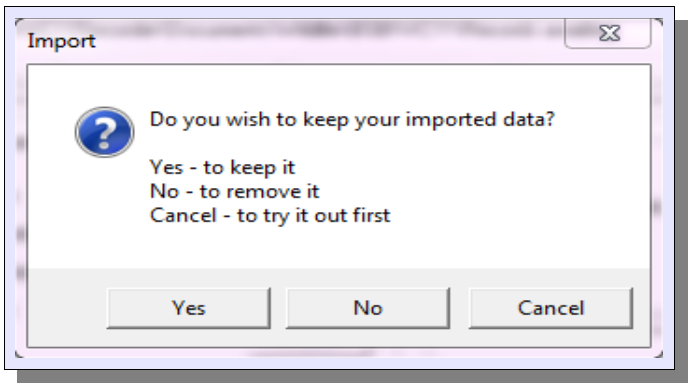


Once you've aligned what you already have on MapMate with what you are about to receive, you are ready to carry out the import. While it's important to fix any problems like bad Taxon names that will stop a Record being loaded at all, remember that any of the other issues can either be dealt with now or

remedied later, once you have the data on board, by using the normal facilities for browsing and editing.

MapMate by default will reject any records with errors, but it's possible to turn this off by unchecking the **Reject records with Errors** box. I advise you not to do this. It will result in your holding records on your MapMate database that have incomplete related data, and will not normally display in reports, atlases and so on. It's possible to bring these to light using the **Browse all Records (ignoring defaults)** report from the **Analysis / Browse records** menu, where they will show up as records with some of the data missing. But they will be mixed in with deleted records (all data missing), and records from outside the currently selected taxonomic and geographic areas of interest, which will **also** appear to have data missing on this report. The only way to check is to bring the individual record on view, in which case the "out of area" ones will appear with all detail intact – and the broken ones will still have it missing. In a large database this is tedious. I prefer not to take on erroneous data in the first place.

Proceed to the **Import Now** button. Depending on the number of records in your file and the speed of your computer, the process can take anywhere from a few seconds for a few records to an hour or more for a hundred thousand or so. Let it run its course – although, if your computer should fail in the middle of an import, it's possible to recover from this, as we shall see shortly. Finally, click on the **Done** button.



This will give you three options. If you've validated and verified all the details of imported records before this stage, you may be happy to select **Yes** and accept the imported data as full members of your database.

If, on the other hand, you realise that there is something seriously wrong with the import, you can click on **No** to discard it. One important thing to appreciate is that, unlike syncs, running the importing process twice on the same file matters. It will generate a duplicate set of Records on your database (but with different Global Unique Keys, obviously). So if you reach this point and realise you have accidentally reprocessed the same file again, it's a good moment to say **No**. The same is true if your computer crashed during a previous import.

Most likely you will want to carry out further review on the imported Records and associated data, now that you have them in a place that makes it easier to check and compare them with your other data holdings. In this case, choose the rather unfortunately named **Cancel**. This gives the imported data a special 'holding' status on the database.

To view just the data from the last import, you can run the report **Browse all Records in 'quarantine'** from the **Analysis / Browse records** menu. This will bring up the usual Data Browser with these records in it. At this stage you can select them for editing. If you want to check them off against other Records you already hold, you can use any of the methods introduced in Chapter 4 to list Records by any chosen criteria; they still appear in these

reports.

When you are ready to impose your final judgement on the imported records, you should return to the importer form via **File / Import / Data from Tab Text Files...** (Make sure you closed it down in the meanwhile!) You will then see a warning **You have Imported data that is not yet saved...** You can click on **OK**, and you will get the same three options as you did at the end of the import. You can't get past this point unless you choose **Yes** or **No**; in other words, you can't accumulate more than one set of imported data in 'quarantine'. If you don't actually have any more text files to import at this time, you can **Cancel** at the next stage.

Export / Import with other MapMate Users

We've discussed earlier the circumstances under which you might want to accept text imports, rather than the generally more desirable sync files, from other MapMate users. Here we look at the mechanics of the process.

Generating an Export Format File

This is pretty straightforward for most purposes; select the Data Entry form, and populate it with some criteria for the records to export. If you're unsure on this, refer back to the section *Simple Ad Hoc Queries: Listing Records* in Chapter 4.

Now select **Query / Records in Export Format...** from the Data Entry form menu. You can make additional manual selections of records from here before using the **Save Selection** button to create the file. This can be imported into another MapMate license without more ado. By default the file is saved in the 'Data / Output' folder of your MapMate setup, but you can change this to suit your circumstances.

No Reference to Reference

One thing you'll notice is that the spreadsheet import process doesn't allow any entry for the Reference item in MapMate. In fact, the importing process automatically generates its own Reference, indicating that the record has been imported by you on a given date from a text file. If you are exporting files from a MapMate source, any Reference that was in the original Record will be lost.

Web-based Data Feeds

It's becoming increasingly popular to provide Web-based ways for people to log their observations and contribute to recording projects and schemes. How the data is then dealt with depends, of course, on the individual project, but there are some general-purpose recording tools that are worth taking notice of, as they may provide you with important data sources in the future.

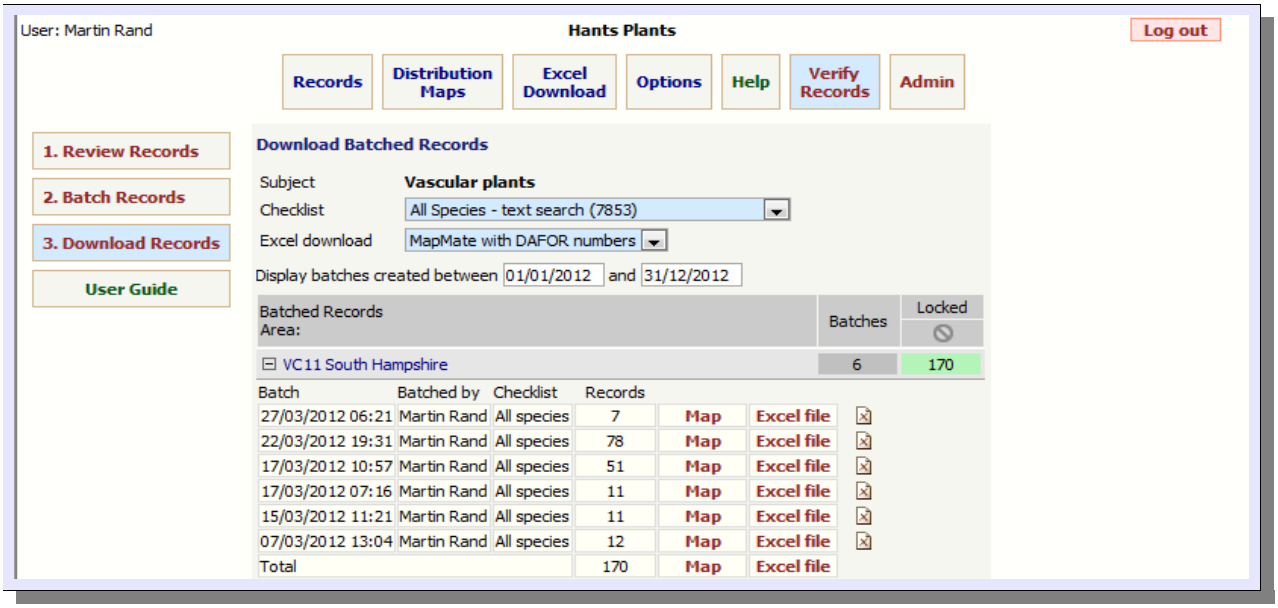
One of these is the *Indicia* project promoted by the National Biodiversity Network. This is not actually a stand-alone recording program like MapMate, but a toolkit that recording scheme promoters can use to develop and customise their own recording Web sites. At the time of writing there is a small number of projects up and running, only one of which involves vascular plants, and then only on a very small scale. Also, *Indicia* doesn't standardise the way in which data is extracted for inclusion in other data stores such as MapMate, so I can't really generalise about this except to point out again the requirements for text file import covered in *Column Contents* earlier in this chapter. If you are taking data from a third-party *Indicia*-based recording scheme into MapMate, you are recommended to discuss these requirements with the developers of the scheme.

One general-purpose Web-based package that has standardised its data export facilities is *Living Record,* developed by **mc² Data Innovation Ltd**. The general Web portal for this is at http://www.livingrecord.net. You may also find it embedded into the Web sites of other recording bodies; for instance, it appears as a page in the Hants Plants Web site http://hantsplants.org.uk which is used to support the activities of the two BSBI vice-county recorders for Hampshire. If you come across it in the latter way, then you will usually find that signing up through the dedicated Web page will make you a member of the local or project-based recording group, and your details will be visible to the person running that group. However you are free to become an 'uncommitted' member by signing up through the first route. Record collators and verifiers for a given region or scheme will still be able to see and make use of any records you enter under their scheme category.

If you are the defined 'verifier' for a given *Living Record* scheme (for instance, I am the verifier for 'Hampshire South VC11') then life is easy. The option to **Verify Records** allows you to go through a three-step process of verifying records, batching them up for export, and downloading them. We won't detail the whole process here, which is well documented on the *Living Record* site. But when you come to the stage of downloading records (see the illustration below), you should make sure to select the option **MapMate with DAFOR numbers** as the **Excel download** format. This will allow you to bring records into MapMate without further ado.

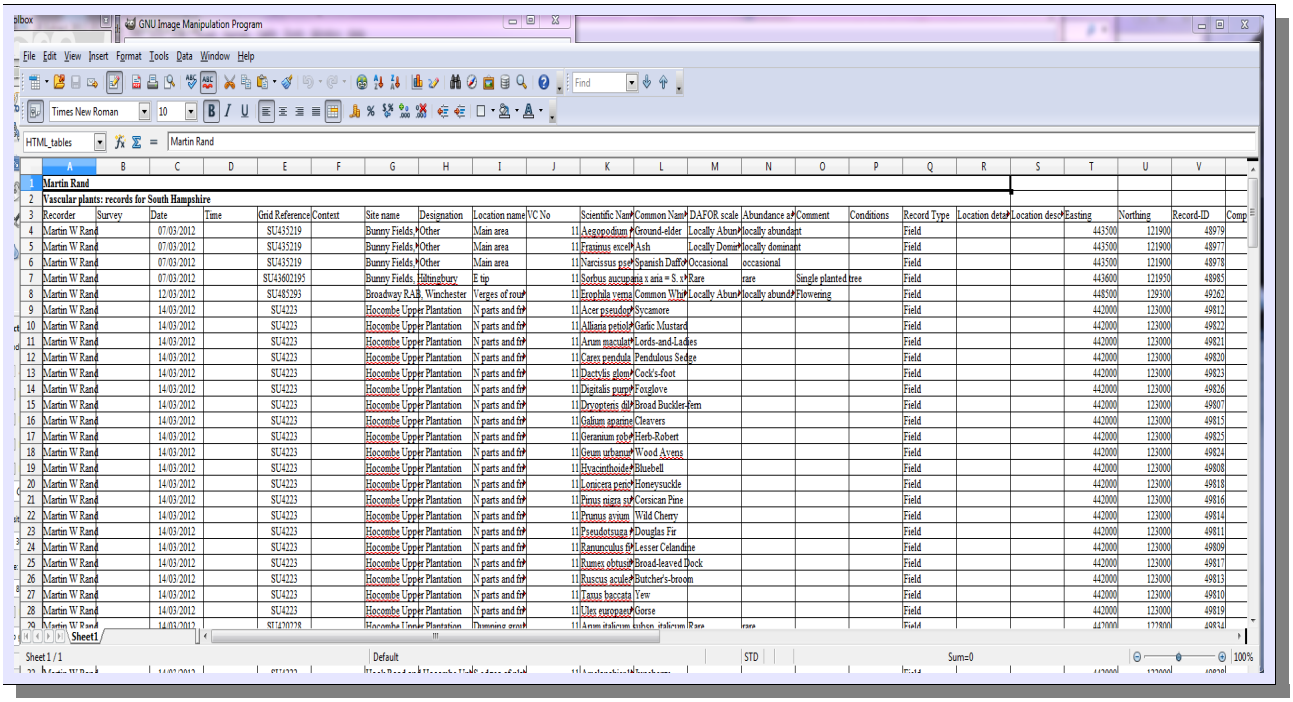
If you are a record submitter and you know that your vice-recorder is a signed-up verifier with *Living Record,* you can be assured that any records you enter can be loaded into their MapMate database – but you will make their job easier if you observe a couple of rules.

- Don't make Site names longer than 64 characters.
- Follow any conventions for Site naming and Specimen referencing that the verifier wants you to observe.



If you are not a verifier, and are not submitting records for a verifier to collect, but have decided that you like the map-based entry facilities provided with *Living Record* and want to use it as your own 'front end' for MapMate, you have a slightly more complicated job to do, as the program doesn't provide a directly MapMate-compatible export format for ordinary recorders.

You need to select the **Excel Download** option from *Living Record*'s main menu bar. This give you the option to **Extract** (your own records only) **to Excel**. Once you have saved the file you can then reload it into either *Excel* or OpenOffice *Calc*.



This clearly isn't the layout and content that MapMate requires for a text import, but all the basic information is there. By renaming and reordering some columns, and removing

others, it's possible to arrive at the required specification. Alternatively, you can use the workbook template available for download at *URL reference goes here*. By copying and pasting the exported data into the first sheet of this template (**Data from LR**), you will automatically get the data you need for writing to the tab text file in the second sheet (**Data to MM**). You will need to delete unused rows from this second sheet, as described earlier for spurious blank lines.

You may also like to consider that at the present stage of development, *Living Record* doesn't give you all the data items you might want to record: full population counts and estimates, Stage and Status for instance. If these are important to you or to your recording partners, MapMate is the better tool to use for data entry.

Validating Data

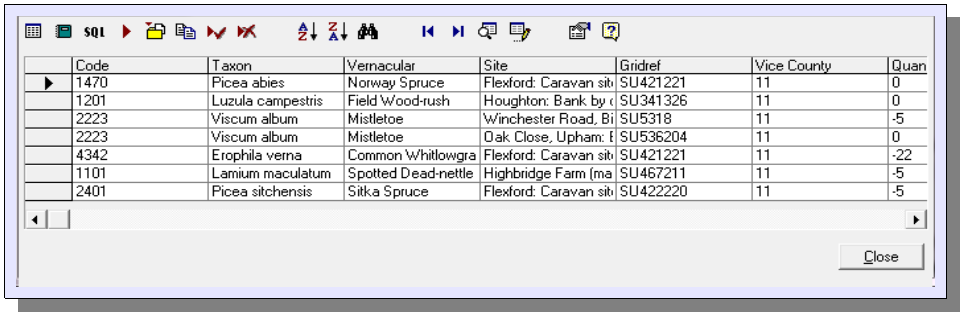
If you are taking in data from partners and contributors to be passed on to others, the correctness of that data will be of concern to you. There are two aspects to this: validation, or the tracking down of data errors, and verification, which is assurance that the plant recorded has been accurately determined, ascribed and, if necessary, documented by additional means. Clearly these two can't be completely disentangled: a transcription error in a scientific name and an inaccurate determination can lead to the same result. We present here a number of tools that will help with validation. Verification is something that relies more heavily on your expertise and local knowledge or that of supporting specialists; but there are means, also discussed below, by which you can pick out the records that need further investigation.

Stock Queries Provided by MapMate

Finding Out What You've Received

We've seen that for text imports there is an easy way to isolate the data you've just received in an import file: put the records into “quarantine” and use the stock query provided to examine them.

For data received by sync file, there isn't an equivalent for this, but you can find out what you've had recently by running the query **Browse all Records new at last sync** from the **Browse records** section of **Analysis**. Depending on how you organise your sync exchanges, this may show you more than just the records that came in most recently on a single sync. It simply gathers up anything that has been added to the database since the last sync was run, including records added locally and text import records.



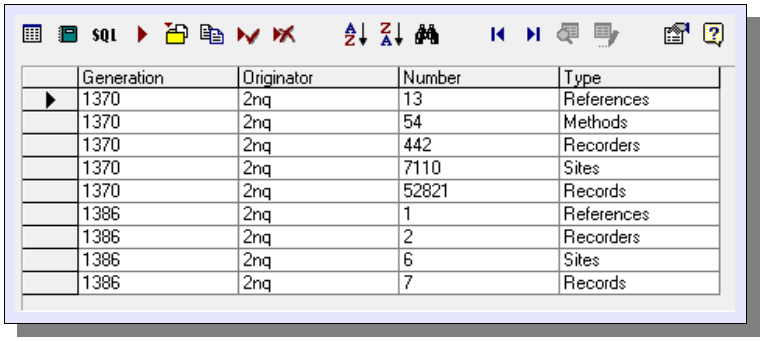
Code	Taxon	Vernacular	Site	Gridref	Vice County	Quan
1470	Picea abies	Norway Spruce	Flexford: Caravan sit	SU421221	11	0
1201	Luzula campestris	Field Wood-rush	Houghton: Bank by	SU341326	11	0
2223	Viscum album	Mistletoe	Winchester Road, Bi	SU5318	11	-5
2223	Viscum album	Mistletoe	Oak Close, Upham: I	SU536204	11	0
4342	Erophila verna	Common Whitlowgra	Flexford: Caravan sit	SU421221	11	-22
1101	Lamium maculatum	Spotted Dead-nettle	Highbridge Farm (ma	SU467211	11	-5
2401	Picea sitchensis	Sitka Spruce	Flexford: Caravan sit	SU422220	11	-5

Close

For instance, this is what I see if I run the query against the MapMate system I'm using to produce the illustrations here, where I don't do syncs with partners. These are records added through a text import.

While this is not quite as neat as the 'quarantine' facility, it usually isn't difficult to sort out the records you're interested in reviewing from this list.

Sometimes one wants to delve back to revisit a batch of data that came in at some point in the past. MapMate doesn't keep any track of 'date record entered' or 'date record acquired' for the individual record, so you won't be able to use that means. However there is a section of the **Analysis** menu, **Database info**, that is of great use here. The starting point is the query **List Database History by generation**.



Generation	Originator	Number	Type
1370	2nq	13	References
1370	2nq	54	Methods
1370	2nq	442	Recorders
1370	2nq	7110	Sites
1370	2nq	52821	Records
1386	2nq	1	References
1386	2nq	2	Recorders
1386	2nq	6	Sites
1386	2nq	7	Records

This shows you how many, of what, you received from each of your sync partners, or entered or imported for yourself, at a particular generation number on your MapMate copy. Don't worry about the fact that generation numbering is not consecutive; it doesn't mean that data has gone missing. (The history here is particularly bizarre, as a lot of early generations are missing, but that's a quirk of the way I set up my MapMate copy for illustrations.)

This still doesn't get you to a range of dates when data was taken on, but you will probably be able to infer from the pattern of records which generation or generations are of interest to you. Now you can run one of the queries such as **List Records introduced at <generation>** to let you see the detail.

Validation and Verification Aids

MapMate has a very limited number of Stock Queries to help with these processes, but it's worth looking at what's available in the **Species 'new or absent'** section of the queries. Here you will find **New species in <year>** and related queries that will allow you to pick up on egregious records. Clearly this is more effective if you already have an established base of good records for your area.

More Useful Queries

For more serious validation you are strongly advised to make use of Martin Harvey's Web site http://sites.google.com/site/kitenetter/Home/mapmate/sql, where you should look for the heading **record verification – set of queries**. Read the notes carefully; if you are uncertain how to copy and paste the text of the queries themselves into your copy of MapMate, which must be done individually for each query presented, refer back to Chapter 4. These queries are mostly of use in highlighting possibly erroneous taxa.

The NBN Data Cleaner

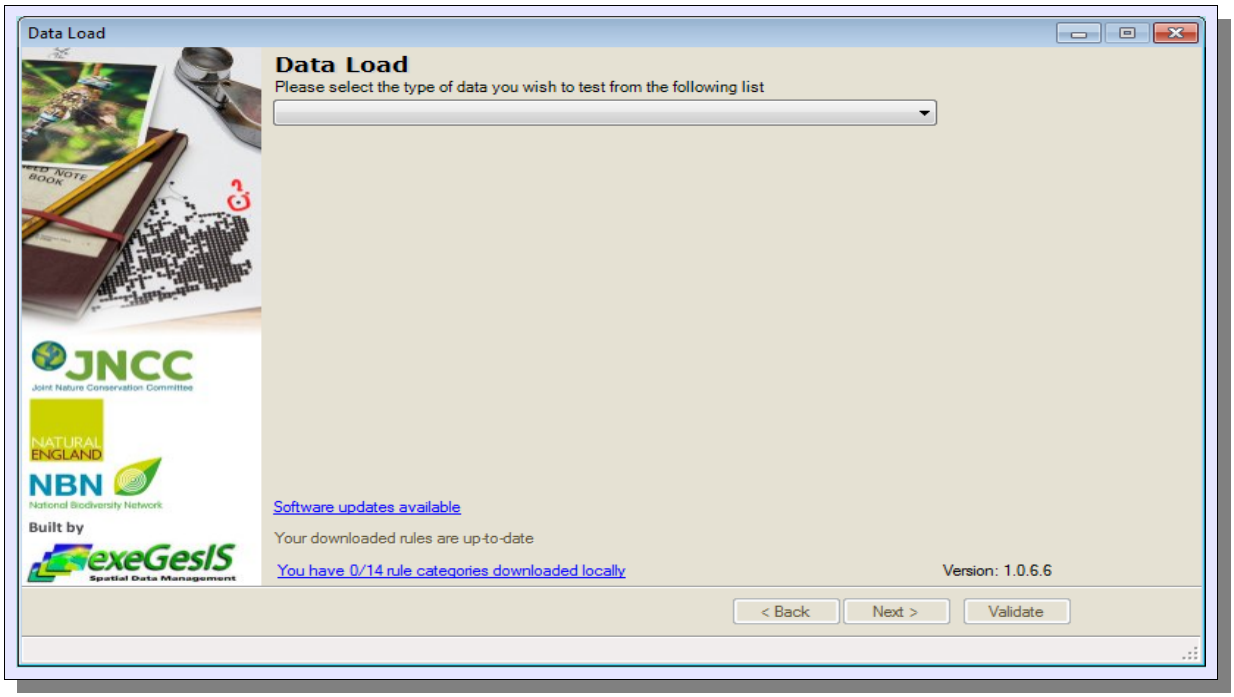
This is a tool made available by the National Biodiversity Network for detecting actual and potential errors in data. It must be run on your own PC, and can be downloaded from the NBN Web site at http://www.nbn.org.uk/Record-Cleaner.aspx.

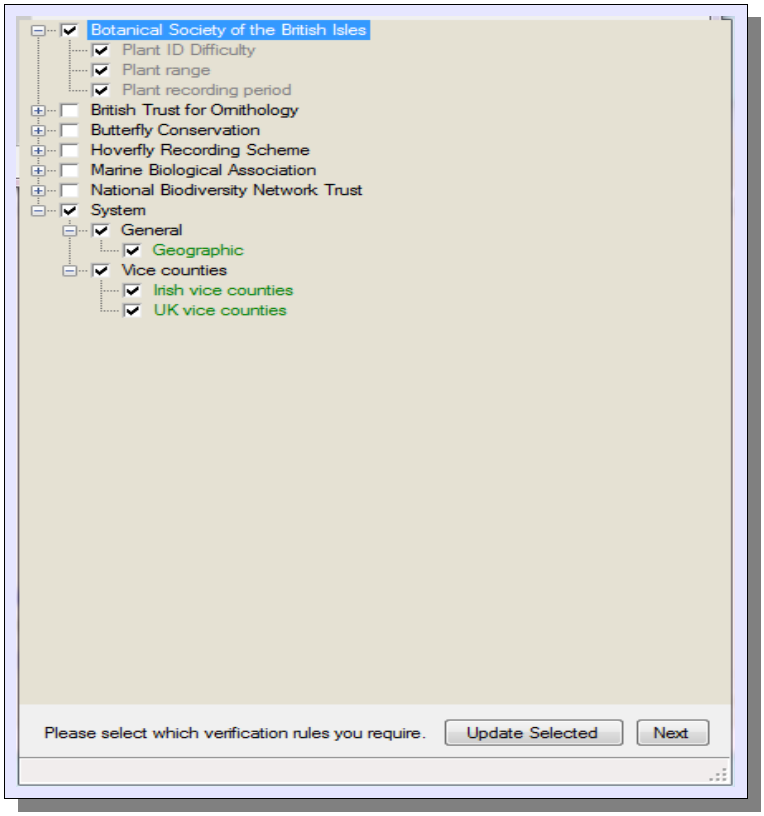
“Cleaner” is a bit of a misnomer, as it doesn't actually change anything on your database; it merely identifies data for attention. You use your normal means for getting corrections to the data, which is as it should be.

I don't intend to present an exhaustive account of the Record Cleaner, as it has a comprehensive Guide that can also be downloaded from the Web site. It can be used on data sets provided in a variety of file formats, and it can also be applied to a MapMate database (when prompted, you need to select the “user.mdb” database file in the Users

folder of your data directory). But unfortunately it fails on large MapMate data sets of more than a few hundred thousand records. If your base of records is already larger than this, your only strategy is to export subsets of the records from MapMate into tab text files and check them one by one.

When you first start the Record Cleaner up, you will see a screen display like the following.



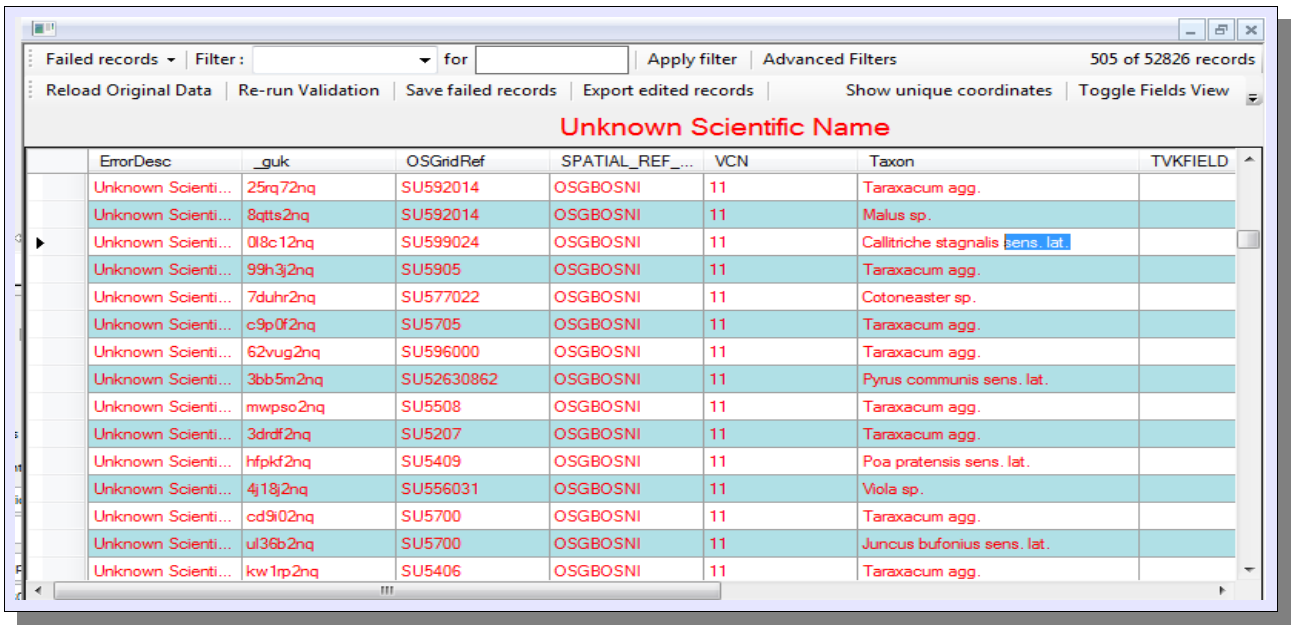


Before you start validation, I recommend that you click on the caption **You have 0/14 rule categories downloaded locally**. Then check the **Botanical Society of the British Isles** rule set and click on **Update Selected**. This will enable you to carry out some verification checks as well as validation. The updating process takes a little while, after which you can proceed with validation.

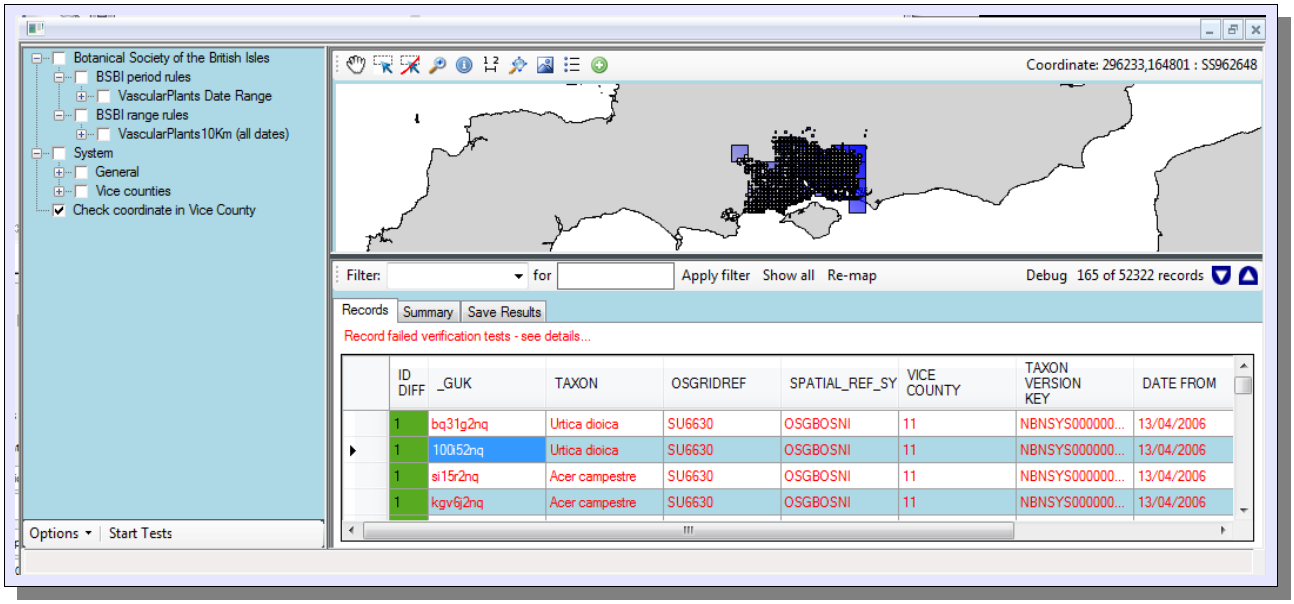
It's also worth noting that you can add your own verification rules, and there is a separate downloadable Guide to this. The process is a bit technical but can be carried out with a text editor.

If you run the BSBI verification rules, be warned that at the time of writing they are incomplete and will generate

lots of range errors against particular species.



Once you have run the validation stage of the process on data, you may end up with errors flagged against one or more Records. You can choose to correct these on the grid display presented to you, so that you can feed these records on to the verification stage, but be aware that you are only correcting them on this grid – not on the source data.



At the verification stage you can choose which tests to apply, and get lists of records that are either questionable or failed tests. The facility to re-map filtered selections of records makes visual checking of locational errors simpler.

Both validation and verification queries can be saved to files for further reading and investigation.

The range of checking that Record Cleaner provides is currently unrivalled, and I recommend that you use it if you are able.