Chapter 7

Conclusions and suggested future work

The retrieval of lake surface temperature observations from ATSR2 data has been investigated and a method of assimilating these observations into a simple lake model has been demonstrated. This method could be used to significantly improve the representation of lakes within numerical weather prediction models. Conclusions about each stage of this process (*i.e.* retrieval, cloud and ice detection, lake modelling and assimilation) and suggestions for future work are summarised in this Chapter.

7.1 Retrieval of LSTs in cloud-free conditions

The application of SST retrieval coefficients to LST retrieval was investigated by radiative transfer modelling (Section 2.2). Validation was also carried out for Lake Tahoe (Section 2.3). Retrieval biases were less than or equal to 0.5 K (Figure 2.4 and Table 2.4) and specific coefficients for LST retrievals were not developed, because errors in cloud masking and ice detection, from using an SST cloud mask for lakes, had larger impacts on retrieved LSTs. Section 1.3 shows that this retrieval accuracy is sufficient to significantly improve the accuracy of LSTs in NWP models and satisfies the ECMWF target accuracy of ± 0.5 K (though they are biases on retrievals from 5 km by 5 km areas rather than errors on a five day mean LST).

The radiative transfer modelling results show that applying SST retrieval coefficients to modelled ATSR2 lake data leads to acceptably small retrieval biases and standard deviations from dual view retrievals. Biases are less than 0.3 K during the day and less than 0.2 K during the night (Figure 2.4), standard deviations are less than 0.3 K during the day and less than 0.1 K during the night (Figure 2.5). The results also suggest there may be a cold bias in dual-2 (daytime) retrievals from lakes at high altitudes ($\simeq$0.2 K for 3000 m to 4500 m). Large ASTDs were the largest source of error in these modelled results. Errors for retrievals using only the nadir view are larger and they are less robust to variations in ASTD, particularly if two channels are used (biases were up to 1 K, Figure 2.6, and standard deviations up to 0.6 K, Figure 2.7). This shows one advantage of using a dual view radiometer, *i.e.* the retrievals are more robust to effects not included in the design of retrieval coefficients, because of the added constraint of sampling geometrically as well as spatially. Modelled dual view retrievals were robust (biases and standard deviations changed by < 0.1 K) to changes in aerosol type and number density for surface visibility > 5 km (Figure 2.11 and 2.12). If various ASTDs were applied to the profiles from which the SST coefficients were derived, it may be possible to produce LST coefficients that are more robust to variations in ASTD. Alternatively a global set of oceanic and continental atmospheric profiles, with various ASTDs applied to the continental profiles, could be used to produce LST coefficients. Neither of these approaches was followed, because developing a cloud-masking scheme for lakes was found to be more important for reducing errors in observed LSTs.

Hook *et al.* (2002) showed a cold bias (0.7 K) in nadir-view only two-channel retrievals from ATSR2 for Lake Tahoe. Re-processing this data (received from JPL, Section 2.3) showed that dual-3 biases were small (0.0 K), but there was a cold bias in dual-2 retrievals (-0.5 K, Table 2.4). Retrieval standard deviations were 0.4 K during the day and night. These standard deviations are consistent with the smaller modelled standard deviations, because (i) satellite data from a 5 km by 5 km area is being compared with a point *in-situ* radiometric measurement and (ii) there is probably a 3 pixel geo-location error in the forward view data obtained from JPL, because of the lake altitude. Horrocks (2002) shows that dual-2 ocean retrievals can be 0.3 K colder than dual-3 ocean retrievals. The bias in dual-2 retrievals from Tahoe may occur for the same reasons, or may be because of the lake altitude. Validating retrievals from other small lakes, at different altitudes, would be very useful in resolving this issue, but the required *in-situ* data were not available (Section 1.5).

7.2 Cloud masking of ATSR2 lake images

Using the ATSR2 Averaged SST product as an input into lake models within an NWP system, would lead to biases in assimilated LSTs and a loss of retrieval frequency, because many SADIST-2 cloud tests are not suitable for lakes (Section 3.1). The standard SADIST-2 cloudmask was adapted for lakes (Chapter 3) and applied to four mid/high-latitude lakes (Chapter 4). Biases resulting from cloud-masking were reduced, particularly for large lakes (by up to 7 K for Great Slave Lake, area $\simeq 20000\,\mathrm{km}^2$, Section 4.1.4). Snow/ice were also masked during the day by the new visible *versus* near-infrared test (Figures 4.6 and 4.7). Biases remained from over-masking of lake edges in spring, which could reach 3.5 K for some retrievals from small lakes (Figure 4.20), and large ASTDs could still be masked as cloud. An alternative method of cloud-masking, that reduced these biases from over-masking was developed in Chapter 6 (for example Figure 6.18). This scheme also reduced over-masking of large ASTDs (Figures 6.19 and 6.21).

The impacts of changing the cloudmask on series of LST retrievals were investigated for four mid/high latitude lakes (Chapter 4). This allowed the biases introduced by the masks to be compared and some cloud contaminated retrievals to be identified. *In-situ* observations of LST, at a variety of locations within a variety of lakes (different sizes and latitudes) would allow cloud-masking to be assessed more quantitatively, because then over-masking of cloud-free pixels and under-masking of cloudy pixels could easily be identified. Obtaining such *in-situ* data is difficult (Section 1.5). Before any lake cloud-masking scheme was applied in tropical areas it would also be important to evaluate the scheme in these regions. In particular, though the new static 1.6 μm *versus* 3.7 μm test is intended to be globally applicable, it is possible that it could mask regions with very low atmospheric transmissions at 3.7 μm, because of higher water vapour loading (Section 3.3.1).

7.2.1 Adaption of the SADIST-2 cloud masking scheme to lakes

Section 3.1 shows how systematic over-masking of warm or cold LSTs can occur if the SADIST-2 cloud tests are applied to lakes. In particular the gross cloud test can mask cold LSTs (Figure 3.2), the 3.7 μm & 11 μm dual view test and fog/low stratus test can mask positive ASTDs (Figures 3.14, 3.16, 3.4, 3.5 and 6.19) and the medium/high cloud test and the cirrus test can sometimes mask negative ASTDs (Figures 3.3, 3.6 and 3.7). The large scale spatial coherence test and thermal histogram test can mask cold LSTs (Figures 3.10 and 3.19) and the small scale spatial coherence tests can mask spatially varying LSTs, particularly warm areas near lake edges (Figure 3.8).

The spatial coherence tests and thermal histogram tests were removed from the scheme because these tests caused the largest biases in retrieved LSTs. Many cloudy pixels were passed by the resultant scheme particularly during the day (since the SADIST-2 reflectance test is not applied to pixels that failed the SADIST-2 spatial coherence tests, Figures 4.1 to 4.4). When the SADIST-2 reflectance test was replaced by the new 1.6 μm *versus* 3.7 μm test (based on Karlsson (1996), Section 3.3.1) and the new 0.65 μm *versus* 1.6 μm test these daytime errors were significantly reduced (Figures 4.2 to 4.4). The 1.6 μm *versus* 3.7 μm test has the advantage that it masks spatially homogeneous low cloud, which is sometimes not masked by the SADIST-2 reflectance test and is otherwise only masked by the large scale spatial coherence test or the thermal histogram test (Figure 4.5). The 0.65 μm *versus* 1.6 μm test (Section 3.3.4) was designed primarily to mask ice/snow, but also masks some clouds, particularly ice clouds. 0.65 μm and 1.6 μm data were also used to detect ice/snow and though this was not thoroughly evaluated, the test appeared to be successful for solar elevations greater than 28° (Figures 4.6 and 4.7). It was also found that the NDSI (Hall *et al.*, 1995; Lawrence *et al.*, Accepted, 2002), which also uses one visible channel and the 1.6 μm channel could sometimes classify cloud as snow/ice.

New standard deviation tests were introduced to attempt to reduce over-masking and remove any remaining cloud contamination. These used local standard deviation in 1.6 μm *versus* local standard deviation in 11 μm during the day (Section 3.3.2) and local standard deviation in 11-3.7 μm at night (Section 3.3.3). To attempt to reduce contamination from low cloud, low thresholds were used and the edges of the cloud mask were expanded. This does mask most remaining cloud, but the over-masking of spatially varying LSTs can be similar to the SADIST-2 mask (Figure 4.20). For this reason the final edge expansion step of the new mask (developed in Chapter 3) should probably be removed and the SADIST-2 small scale spatial coherence test should possibly used instead of the new night-time standard deviation test at night (because it can sometimes mask more low cloud). This was, however, not evaluated, because results from the large and small scale spatial coherence tests are combined in the SADIST-2 cloud-mask product.

The cloudmask developed in Chapter 3 can sometimes lead to a loss of retrieval frequency for Lake Vattern, which is approximately 20 km by 90 km (Section 4.1.4). This suggests that this scheme could not be applied to lakes with diameters less than 20 km, though removing the cloud-edge expansion step (as described above) should make the scheme applicable to slightly smaller lakes. Such lakes currently occupy approximately 25% of the land surface in an ECMWF grid box and this just meets the target diameter of 20 km suggested by ECMWF (Section 1.3).

7.2.2 The spectral cloud masking scheme

The success of the new spectral tests, compared with the standard deviation tests, motivated the work described in Chapter 6, *i.e.* the development of a cloud-masking scheme that used all thermal BT information simultaneously and only used spatial tests at cloud edges. The scheme compares observed BTs with a set of modelled BTs, calculated from a global set of marine atmospheric profiles, with various ASTDs applied (Section 6.2.3). The thermal infrared test in this scheme assumes that the probability that a pixel is clear is related to its normalised distance from the pixel to the closest modelled point in this BT phase space. The unnormalised distance in each component of this space was also used. These distances were thresholded to give a clear/cloudy test (for example Figure 6.3).

In addition to the two static threshold reflectance tests developed in Chapter 3 (Sections 3.3.1 and 3.3.4) and applied in Chapter 4, two dynamic reflectance tests were introduced. The dynamic 1.6 μm *versus* 3.7 μm test successfully masked some clouds not masked by the static threshold test (Figure 6.4). Large numbers of pixels with sub-pixel cloud contamination could result in the threshold being set too low and this test could be improved by relaxing the threshold for larger 3.7 μm TOARs (Section 6.5.1). The dynamic 0.65 μm *versus* 1.6 μm dynamic reflectance test should be removed from the scheme developed in Chapter 6, because it can mask turbid water as cloud (Figure 6.16) and if land pixels are used in the threshold determination it can mask water as cloud (Figure 6.6). The values thresholded in the 1.6 μm *versus* 3.7 μm tests were also combined with the normalised distance from the thermal infrared test, so if a pixel just passed both the thermal and reflectance tests it could still fail when they were combined.

There are probably similar amounts of cloud contamination using the mask developed in Chapter 6 to that of Chapters 3 and 4 (Figures 6.8 to 6.11). Large ASTDs were, however, not masked as cloud by the mask developed in Chapter 6 (Figures 6.19 and 6.21) and systematic masking of spatial variations in LST was reduced (Figure 6.18). This reduction in over-masking means that the cloud mask developed in Chapter 6 can be applied to smaller lakes than the cloud mask developed in Chapter 3. Figure 6.18 suggests that it cannot be applied to lakes with diameters less than 8 km. Such lakes currently occupy approximately 5% of the land surface in an ECMWF grid box, and are smaller than the target diameter of 20 km suggested by ECMWF (Section 1.3).

Despite using all infrared BTs simultaneously in Chapter 6, low water clouds could still pass all cloud tests at night (Section 6.4). Figure 6.25 shows that this is not because the test thresholds were set incorrectly, but because low cloud and cloud-free water can have sufficiently similar BTs that infrared tests cannot in principle discriminate between them purely on spectral signature, other information must be used. Spatial tests are a possibility, but low water clouds often form in spatially homogeneous layers and so are not necessarily distinct from water. *A priori* knowledge of the LSTs could also be used, but this will only be effective if the low cloud decreases the retrieved LST to a value outside of the range allowed by the modelled LST. There can be a wide range in LSTs over a lake in spring (for example Figure 6.18 shows a range of 12 K for Lake Vänern in May), so if cloud is present over a warm lake area it may not be detected if a single modelled LST is used. Using more than one one-dimensional model per lake, using a three dimensional lake model, or maintaining a map of variations in LST from the mean LST, would all help to solve this problem. Any cloud masking scheme using a modelled LST should include in the model variance a term to allow for processes not well represented in the model, for example diurnal variations or cold upwellings (Figures 6.12 to 6.15), otherwise cold cloud-free water may be masked as cloudy.

The success of this spectral cloud-mask suggests that it would be worthwhile replacing the reflectance tests used with tests similar to the thermal infrared "distance from model" tests, applied in Chapter 6. A radiative transfer model could be used to generate a global set of reflectance values for the 1.6 μm and 3.7 μm channels, in addition to thermal infrared data. These tests would then not only compare TOARs in different channels, but also include the geometric constraints on the TOARs in each view. Variations of TOARs with low sun angles would also be accounted for and all BTs and TOARS would be used simultaneously. It is possible that the dynamic 1.6 μm *versus* 3.7 μm test would still mask pixels passed by these RTM based tests, because the dynamic test compares a pixel's TOARs with other TOARs from the same image. A comparison of visible and 1.6 μm TOARs is needed to mask snow and ice during the day, but modelling visible channels is more difficult than modelling near-infrared channels, because they can be affected by water turbidity (Figure 6.16). So, a static threshold snow/ice/cloud mask should be applied before the RTM based cloud mask.

The spectral algorithm applied in Chapter 6 is slow (5 minutes for a large lake, with the code written in IDL, which was not designed for computational speed).

Instead of calculating the distance to the model calculating the means, variances and covariances of the modelled BTs and TOARs in each bin and using these to calculate the probability that the pixel is clear would improve the speed.

7.3 Assimilation of satellite observations into a one-dimensional lake model

Assimilating ASTR2 LST observations into a one dimensional lake model allowed model errors that occurred on a time-scale greater than one week to be corrected, whilst the model still accounts for variations in LST that occur on shorter time-scales (Figures 5.14 to 5.17).

Modelled LSTs, with no observational input, differed from LSTs driven by assimilated observations by between -3.7 K to 5.5 K for the four lakes investigated (Section 5.5.2). Differences between the ECMWF lake climatologies and the assimilated LSTs were much larger (-15 K to 2.0 K, Section 5.5.2). There were significant differences between modelled, assimilated and ECMWF climatological freeze/thaw dates; in particular the ECMWF thaw dates and the observed thaw dates differed by up to 20 days and Lake Vänern did not freeze in the winter of 1999/2000, whilst it was frozen for 192 days in the ECMWF climatology. These differences between the ECMWF LST and freeze/thaw climatologies and the assimilated values should have significant impacts on modelled meteorology near the lakes (Ljungemyr *et al.*, 1996). The combination of lake modelling and data assimilation represents variations in LST that occur on shorter time-scales than the target of a five day mean suggested by ECMWF (Section 1.3), but the combination of retrieval biases, remaining biases from over-masking (particularly in spring), or the divergence of the model from reality during cloudy periods may sometimes lead to errors greater than ± 0.5 K, the target accuracy suggested by ECMWF.

The assimilation heat flux scheme used was based on Gustafsson *et al.* (1997), but the assimilation flux was applied at the surface, instead of the top 10 m of the water column. Most of the real heat fluxes between the lake and the atmosphere occur at the surface, so the assimilation flux can be viewed as a correction to those fluxes. The surface assimilation heat flux often affected a layer either less than, or much more than, 10 m deep, depending on how well heat was mixed downwards from the surface, which is largely controlled by the depth and strength of the thermocline (Figure 5.12 and 5.13). In these cases there will certainly be significant differences between results from the surface flux scheme and the method used in Gustafsson *et al.* (1997).

Measured lake temperature profiles were not available, so the effects of the assimilation heat flux on temperatures below the surface could not be validated. However, effects of the assimilation flux on the surface temperature and on LST forecasts were evaluated (Sections 5.4 and 5.5.3). The difference between the assimilated LST (*i.e.* the LST at the end of the assimilation period) and the analysis LST (*i.e.* the best estimate of LST at the end of the assimilation period, made from the modelled LST and the observations) was evaluated for four assimilation methods. This difference was typically 0.1 K$\pm$0.4 K for the best method, which based the assimilation flux on both the estimated wind mixed depth and the estimated thermocline depth, as well as the difference between the modelled LST (*i.e.* the LST without data assimilation) and the analysis LST. This method still tends to under-assimilate for deep cold mixed layers. Section 5.5.1 suggests that the assimilation flux should be less windspeed-dependent for such layers. There is also some over-assimilation for shallow warm mixed layers. The assimilation flux only affects a layer 5-10 m deep for many of these cases and so improving the model resolution (2.8) near the surface may reduce this over-assimilation. Measured lake temperature profiles would be useful in validating this assimilation scheme, or a variational scheme based on this surface flux scheme (Section 5.6).

The lake model used a very simple ice model, but an improved version is available from SMHI. This is probably the single biggest improvement that could be made to the lake model (Section 5.2). Observations of lake ice were also not assimilated. So, the assimilation heat flux forced the modelled LST towards the temperature of any ice-free portions of the lake, even when the lake was mainly ice covered. This leads to errors near freeze-up/melting, which could be avoided if ice cover was assimilated. Night-time retrievals should not be assimilated when the lake may be partially, or wholly, ice covered, because ice may not be masked without using reflectance tests.

Using separate models for different areas of the lake would allow partial ice cover and spatial variations in LST to be represented. This could also be used to improve cloud masking, by providing better *a priori* LSTs. It would be interesting to investigate whether data assimilation can adequately account for heat exchange between the different models. The lakes could be split by topography/depth (as applied by Gustafsson *et al.* (1997) to the Baltic Sea), or in a regular grid. Splitting the lakes using a grid makes modelling the exchange between the lake and the atmosphere in an NWP model simpler, but there will be larger variations in LST within each lake grid-box than if the lake is split by topography/depth.

7.4 Future application to numerical weather prediction models

There are advantages to using both lake models and LST observations within an NWP model. Using a lake model without assimilating observations can lead to significant errors in both LST and in melt/freeze dates (Figures 5.2 to 5.5, Table 5.4 and published results described in Section 1.3). These errors will have a significant impact on forecast meteorology (Ljungemyr *et al.*, 1996). Using a lake model does, however, allow variations in LST and ice cover, occcurring on time-scales smaller than the inter-observation period, to be accounted for (for example there is an LST change of 7 K in 10 days in Figure 6.11) and LST forecasts to be made. Lake models can be constrained by assimilating ATSR2 observations (Chapter 2 shows that we expect retrieval biases in cloud-free areas to be small, *i.e.* less than 0.5 K). Chapter 5 develops one possible assimilation method.

Assimilating the ATSR2 GSST product into a lake model can lead to significant biases in springtime LSTs for mid/high-latitude lakes (Figures 4.27 to 4.30), because of the application of the SADIST-2 ocean cloud mask to lakes. In addition the SADIST-2 scheme does not detect lake ice and so ice cover cannot be assimilated and LST retrievals are made from ice covered pixels. Using either the SADIST-2 based mask developed in Chapter 3, or the spectral mask developed in Chapter 6, avoids these problems (Figures 6.8 to 6.11). However, over-masking using the mask developed in Chapter 3 can lead to a loss of retrieval frequency from smaller lakes (diameters less than $\simeq 100\,\text{km}$). This problem is significantly reduced if the spectral mask is used. The spectral mask also reduces over-masking of large magnitude ASTDs (Figures 6.19 and 6.21), though this seems to have little impact on the resultant assimilated LSTs. It seems likely that a quicker algorithm could be developed, that achieves similar results, without an unacceptable increase in cloud contamination.

In the future a useful global LST product could be produced from Advanced ATSR data, in addition to the current Advanced ATSR SST product. This could be assimilated into lake models within NWP systems to improve the representation of lakes in forecasting models. Replacing modelled, or climatological, LSTs within a current NWP model with assimilated LSTs from ATSR2, would allow the effects on modelled meteorology to be evaluated, with major impacts likely in heavily-covered mid/high-latitude regions (Canada, Scandinavia and north-west Russia) and local impacts near isolated lakes (for example Africa). A similar approach to that applied here to ATSR-series sensors will be necessary to achieve similar products from other candidate observing systems, for example MODIS, AVHRR (these single view instruments will, however, be prone to retrieval biases). This work has significantly addressed the scientific/methodological problems involved in such observations - the groundwork has been done for an operational implementation.