Final report of the
Joint WCRP/SCOR Working Group
on Air-Sea Fluxes
(SCOR Working Group 110)

INTERCOMPARISON AND VALIDATION OF OCEAN-ATMOSPHERE ENERGY FLUX FIELDS

by

Members of the WGASF

Edited by P.K.Taylor

June 2000

EXECUTIVE SUMMARY

The Working Group

The main aims of the Joint JSC/SCOR Working Group on Air Sea Fluxes[1] were to review the requirements of different scientific disciplines for surface flux data sets, to catalogue available surface flux data and flux-related data sets, and to review the strengths and weaknesses of each. The three meetings of the Working Group were (1) NOAA Headquarters (Silver Spring, Maryland, USA), 22nd to 25th October 1997; (2) KNMI (De Bilt, The Netherlands) 14th to 17th April 1999; (3) SOC (Southampton, UK) 6th to 10th December, 1999. The Working Group's findings are contained in this report, which has been prepared on an accelerated schedule at the request of the JSC.

The Fluxes Considered

The fluxes considered by the Working Group were the transfers of heat, water and momentum between the sea and the air. The flux products considered were those which: were readily available to all scientists; were a gridded, derived surface flux or flux related field; covered major ocean basin scales and upwards; and, were adequately documented. The Working Group has assembled a catalogue of flux products which meet these criteria. It is available on the Internet[2] and described in an appendix to this report. Given the limited time and resources, it only proved practicable to evaluate a subset of the more recently available flux products. The Working Group also summarised the State-of-the-Art with regard to the basic meteorological variables (used in flux calculation), and other important components of the climate system such as waves and sea ice.

Requirements and Data Sources

The Working Group has reviewed the requirements for flux fields and the characteristics of the various data sources available for flux estimation. The latter include *in situ* data from ships, buoys, rigs, etc.; instruments on satellites either in polar (or other precessing) orbits or geostationary; and models, run either in operational Numerical Weather Prediction (NWP) or reanalysis modes. The requirements for surface flux fields are many and diverse. However the Working Group's review suggests that there are four main classes of requirements.

One class of requirements is the need for surface flux fields on high time and space resolution; typically 3 hours and 50 km. Such data are needed, for example, for forcing ocean general circulation models, for wave modelling and forecasting, and for regional weather nowcasting and prediction. Such high space/time resolution is only possible at present using models, however in the future, high resolution satellite products will become available. The quality of the wind fields is particularly important for these applications and operational satellite-borne scatterometers are expected to make an important contribution.

A second class of requirements is for flux fields on longer space and time scales but with high absolute accuracy, typically within a few W/m^2. Customers for these data include climate and sea-ice modellers. At present such needs are difficult to satisfy because of the problem of quantifying the systematic errors in the flux data. Only by comparing flux values obtained using different data sources and estimation methods are these stringent accuracy requirements likely to be obtained.

---

1 (SCOR Working Group 110: Intercomparison and Validation of Ocean-Atmosphere Energy Flux Fields)

2 available from the WG web site: http://www-pcmdi.llnl.gov/airseawg/ (mirrored at http://www.lmd.jussieu.fr/pcmdi-mirror/airseawg/ and at http://www.bom.gov.au/bmrc/clch/pcmdi-mirror/airseawg/)

Climate variability studies define a third class of requirements. Again high absolute accuracy is desirable, however consistency and continuity of the data over a suitable time period (and spatial domain) is the vital need. The longest time data set available is from the merchant ship observations (as assembled in the Comprehensive Ocean-Atmosphere Data Set, COADS). However the ship data suffers from poor sampling in many regions and time dependent biases due to changing ship characteristics and observational technique. Comparison with the reanalyses suggests that the COADS data can give useful information on interannual variability of monthly mean fluxes in the mid-latitude North Pacific and North Atlantic, and in parts of the South Atlantic; over much of the tropics and the Southern Hemisphere, north of 40S, the COADS data can only define a long-term monthly mean climatology. However, the reanalyses also suffer from time dependent biases because of the assimilation of data from the evolving observational system. Thus, whether using models or observations, the detection of long-term trends in the surface fluxes is difficult if not impossible. Satellite data has now been available for a significant and lengthening time period. Such data has the potential to provide a consistent time record but only if great care is taken to inter-calibrate sensors on successive satellites.

A final flux requirement is for high quality verification data. For example, NWP models need independent estimates of the basic meteorological variables and the fluxes for verification of the model physics. Such estimates must be associated with a realistic error assessment and specification of the true resolution of the data. Similarly Ocean General Circulation Model development would be aided by air-sea interaction experiments with measurements of the fluxes in both atmosphere and ocean, with ocean mixed layer measurements sufficient to allow budget closure. Such high quality measurements are not easy. The "flux buoy" systems are an example of instrumentation capable of providing verification data for both models and satellite - borne instruments.

Flux Variability

For many applications it is important that a flux product contains accurate information on the variability of the fluxes. In some ways the determination of flux variability is easier than estimating mean flux values; systematic errors are of less importance, and the main spatial scales of longer term variability tend to be large. Numerical weather models tend to perform better with regard to determining the variability than they do in determining the mean; probably because they are targeted at forecasting the weather - the day to day variations in climate.

However despite these considerations, and the undoubted importance of surface flux variability, the Working Group considered that a detailed assessment of variability in the flux products could not be accomplished at this time. The report contains a brief discussion of flux variability to set the context for the flux field evaluations. It was noted that most previous studies had concentrated on the variability of basic variables, such as sea surface temperature (SST), rather than the fluxes themselves, and that further studies focused on the fluxes were desirable. The Working Group concluded that a thorough assessment of the variability of available flux products must remain an important priority for further studies, perhaps by a future working group.

Direct Flux Measurement

Direct measurements of the air-sea fluxes are too few to contribute directly to the calculation of large scale flux fields. Rather they are important for developing, calibrating, and verifying the parameterisation formula used to estimate the fluxes from the basic variables. Thus the accuracy of the direct flux determinations represents an accuracy limit for the indirect estimates. The methods of direct measurement of the radiative fluxes, turbulent fluxes, and precipitation were reviewed.

Accurate measurement of surface radiative fluxes is not easy, and with regard to the instrumentation needed there are contrasting views. Experts in radiation measurement assert that only sophisticated instruments and procedures, as defined by the Baseline Surface Radiation Network standards, will provide adequate accuracy. However such measurements are difficult to implement at sea and will be restricted to very few sites. In contrast, scientists involved in air-sea interaction studies believe that, by avoiding known pitfalls, consistent results can be obtained from simpler instrumentation. Much depends on the accuracy and time resolution sought. The simpler instrumentation should be adequate to highlight the biases in the radiative fluxes in many models (presently a few 10's Wm^{-2} in some regions).

Estimation of the turbulent fluxes using fast response instrumentation mounted on research ships or buoys is becoming more common. The techniques used in the eddy correlation method and the less direct, but more easily implemented, inertial dissipation method are summarised. When implemented on a ship both methods are prone to errors but of differing nature. Thus in future it is desirable that both methods be used to fulfil an important requirement, the determination of the transfer coefficients for flux parameterisation (see below).

Rainfall, particularly that from tropical convection, is highly sporadic both in time and space. For most applications area averaged values are required. These can be obtained by ship or aircraft mounted radars on the local scale, or from satellites in the longer term and on larger scales. However the sensors and retrieval algorithms need to be verified by reliable ground-truth measurements. Accurate measurement of rainfall at sea presents difficulties in two respects; there are measurement (instrumental) problems, and sampling problems. New rain-gauges have been developed for high wind conditions but air-flow distortion by the ship remains a problem. Even during intensive measurement campaigns the spatial and temporal sampling may be inadequate. Results of different satellite algorithm intercomparison projects suggest that presently available *in situ* measurements and radar estimates are not representative enough to serve as a verification data source.

Parameterisation of the Fluxes

Radiative fluxes

The parameterisation formulae for obtaining the net surface shortwave and longwave fluxes from ship observations are relatively crude, relying on the estimate of cloud amount to characterise the effects of cloud on the fluxes. Satellite based estimates use measurements of the top-of-the-atmosphere radiation, and radiative transfer models (RTM's) to estimate the surface value. For clear-sky conditions, the surface LW values from a sophisticated RTM with *in situ* radiosonde data can be used to check the calibration of pyrgeometers. However under cloudy conditions longwave RTM's are limited by the information available on both cloud base height and cloud emissivity. This limits the reliability of satellite estimates of surface longwave; new techniques are being developed but still need evaluation. Knowledge of atmospheric scattering is essential for shortwave RTM's under both clear and cloudy conditions. Uncertainty in the diffuse component may cause RTM's to over estimate the surface shortwave compared to measurements, possibly by a few percent. Despite this, satellite estimates of shortwave radiation are more reliable than for longwave. Atmospheric models (NWP and reanalyses) use simplified RTM schemes for computational efficiency. However at present the limiting factor in most models is the representation of clouds and their radiative effects. In particular, low level stratiform clouds are often poorly modelled.

Turbulent fluxes

The parameterisations for the turbulent fluxes are based on values of basic meteorological variables such as wind speed, temperatures, and humidity. Since these "bulk formulae" are used for flux estimation from not only *in situ* data, but also satellites, and most models, they were reviewed at some length by the Working Group. Despite years of research there is still uncertainty with regard to the behaviour of the various transfer coefficients, in particular for sensible and latent heat at wind speeds over 10 m/s. For this reason, some newer parameterisations use wind stress estimates and surface renewal theory to predict the heat fluxes. However, the variation of the drag coefficient with sea state remains controversial, particularly for sudden storms, or in light wind conditions when swell dominates over the wind sea. The newer algorithms also incorporate treatments of other physical processes, particularly light wind phenomena such as the diurnal warm surface layer and wind gustiness.

Error Estimation

An important requirement of flux products, which is rarely satisfied, is the realistic estimate of accuracy. The error estimates should consider as many contributing factors as possible, both random and systematic. Representative error estimates may be difficult to determine. The report reviews some theoretical approaches, gives practical examples of quantifying random, systematic, and sampling errors, and shows that it is possible to estimate the cumulative errors.

Many applications need gridded flux fields that are complete both in space, and over some significant time period. The non-uniform sampling of in-situ and remotely-sensed fluxes necessitates the use of an analysis method to interpolate and, sometimes, extrapolate the data. Such techniques alter the error budgets of the flux fields. For example, depending on the choice of method and averaging period, sampling errors in remotely-sensed data may dominate the systematic biases. The advantages and disadvantages of two mapping techniques were discussed as examples. The successive correction technique is computationally simple and has been widely used. However it creates fields with spatially varying resolution, and spreads information so as to degrade the flux values in well sampled areas. In addition, error estimates are difficult to determine. Objective analysis techniques which take proper consideration of observational and sampling errors, and are capable of producing a posteriori error estimates, should be employed instead of the sucessive correction method. In particular, state-of-the-art variational techniques developed in the context of atmospheric and oceanic data assimilation should be applied to the analysis of surface marine fluxes.

Evaluation of Available Products

Methods of evaluation

The methods of evaluation for the basic variables and the flux fields include comparisons to reference data sets (such as direct flux measurements or those from high-quality sources) and intercomparisons between products. For the fluxes it is also possible to use integral constraints such as the global heat and water balance, oceanic transports, and enclosed basin budgets, as well as comparisons with other indirect flux estimates.

Basic meteorological variables

It is important to know the accuracy of our estimates of the basic meteorological variables. Errors in these variables affect the accuracy of the surface flux values. Using the basic variables to verify model and satellite derived estimates avoids uncertainty due to the use of different flux parameterisation formula. As a rough guide, achieving 10 Wm^{-2} accuracy in a heat flux estimate requires knowledge of wind speed to a few tenths m/s, air and sea temperatures to about 0.2K, humidity to about 0.3 g/kg; these requirements vary regionally being generally more stringent in the tropics.

For most of the basic variables, the Working Group reviewed the accuracy of high quality *in situ* measurements, standard *in situ* measurements, and satellite based estimates. The high quality measurements, as might be obtained from research experiments and "flux buoys" are of value for verification and for parameterisation development. The standard *in situ* data, for example from Voluntary Observing Ships (VOS), represent our only near surface measurement of many quantities for use both in flux estimation and model assimilation schemes. As instrumentation and retrieval algorithms improve, the satellite remote sensed data are increasingly being used to define the spatial and temporal variations of a given variable while the *in situ* data are used for verifying the retrievals.

The ocean affects the surface fluxes through the sea surface temperature (SST). *In situ* observations of "SST" are often made at a few metres depth; the SST difference due to the cool ocean surface skin can be allowed for, but significant errors may occur if a near surface diurnal thermocline is present. Satellite estimates are limited by the effects of cloud and biases due to cloud contamination (particularly at night) and the effects of aerosols. The better SST analyses use combinations of satellite and *in situ* data. Typical RMS differences between analyses in most regions are a few tenths K, but greater than that in the Southern Ocean, in coastal regions, and in sea ice areas.

Variations of sea surface salinity (SSS) are not significant for air-sea flux calculation (although its mean effect must be allowed for in latent heat flux calculation). However SSS is important for flux verification and for the freshwater forcing of Ocean General Circulation Models. *In situ* salinity data remain relatively few but will increase significantly due to projects such as ARGO. The potential for, and difficulty of, salinity measurement from space was reviewed. Useful data may be obtained in regions of large salinity gradients, however the 0.1 psu accuracy needed by ocean models is unlikely to be achieved by remote sensing in the near future.

Near surface air temperature is not easy to measure accurately in the marine environment. Biases due to radiative heating of the sensor (or its surroundings), sea spray, etc., are common. VOS data tends to be biased high, probably by a few tenths K on average, but by several degrees on occasion. New techniques for estimating air temperature from satellite data have been suggested; further evaluation is needed. Fortunately, except in high latitudes, the sensible heat flux is a minor part of the heat balance. However air temperature is also needed in all regions for stability corrections during latent heat flux calculation.

The latent heat flux depends on the humidity of the near surface air. Humidity measurement on ships is in some ways easier than temperature determination. However it is likely that VOS data are biased high, particularly where fixed thermometer screens are used. Methods of determining near surface humidity data from satellites have improved significantly and useful accuracy is obtained in many regions. However developing an algorithm which performs well both within and outside of tropical regions remains a problem.

There is no accurate absolute standard for near surface wind measurement over the open ocean. Anemometer-based and visual estimates from ships, and buoy measurements, all are subject to systematic and random errors. "Beaufort scale" values for the visually estimated winds have been calculated which, in the mean, result in good agreement between visual and anemometer winds. Importantly, the main reasons for the differences from previous Beaufort Scales are now understood. Buoy winds may be biased low in high sea state conditions. Satellite scatterometer winds are considered significantly superior to ship winds but the algorithms were developed using buoy data so biases may still remain. Passive microwave derived wind speeds are useful for heat flux estimation. They have the potential to provide good sampling and an increasingly long time series.

Waves are an important climatic variable in their own right and may, to some still uncertain extent, modulate the surface fluxes. The amount of wave data from buoys is increasing but the spatial distribution is poor. Wave data from ship reports is generally considered to be of poor quality; however it has been used to construct wave climatologies which are still much in use. One problem is the separate reporting of, but unclear distinction between, sea and swell. Satellite altimeters provide verified global estimates of wave height; algorithms for scatterometer wave estimation are still being developed. Wave modellers believe that, given an accurate wind field of high temporal and spatial resolution (say 3 hours and 50 km), then a good quality wave field can be calculated.

Cloud amount is used in some radiative flux parameterisations and is also important for the verification of NWP and climate models. Sources of cloud data reviewed included ship reports and the International Cloud Climatology Project (ISCCP). Precipitation also may be estimated from ship weather reports; however variable satellite based estimates offer better future potential. The Global Precipitation Climatology Project (GPCP) and the CPC Merged Analysis of Precipitation (CMAP) have produced precipitation climatologies which merge *in situ* and satellite estimates.

River inflow must not be neglected as a component of the ocean freshwater budget. Data are available through the International Hydrography Program (IHP) which published a Global River Discharge Catalogue, and from the Global Runoff Data Centre (GRDC).

Sea ice is an important climate variable which significantly modifies the surface fluxes and is also a component of the ocean fresh water budget. At present, sea-ice models can reproduce a realistic annual cycle of the global sea-ice cover. Probably this is despite large but opposite errors in the fluxes over sea-ice. Now that multidecadal satellite and submarine records of ice extent and ice thickness are available, the challenge is to understand interannual variability. This will require a higher level of accuracy in estimates of the radiative forcing of the ice surface and data on the interannual variability of snowfall. The errors in reanalysis and satellite-derived surface flux fields must be better characterised and then reduced by more thorough comparisons amongst themselves and with *in situ* observations. Boundary layer formulations need to be examined for their applicability in predominantly stable polar conditions.

Flux field products

Newer products based on *in situ* data[3] have benefited from the valuable new data sets constructed by the COADS project. For wind stress the newer climatologies are definitely to be preferred since previous climatologies, such as Hellerman and Rosenstein (1983) or Oberhuber (1988), over estimated the wind stress in most areas. However poor *in situ* sampling in the southern hemisphere, and particularly in the Southern Ocean (where the winds are underestimated in almost all *in situ* climatologies), suggests that the lengthening satellite derived wind data-set is a promising alternative. For the heat fluxes, our inability to balance the long term mean global heat budget together with the uncertainty in the transfer coefficient values, remains a major hindrance to evaluating many of the flux products.

To illustrate the problem: the global heat balance of the SOC climatology was $30Wm^{-2}$; almost exactly the same as the for the UWM product. This was despite the use of corrections for observation biases in the *in situ* data. In a tuned version of UWM, shortwave radiation was decreased (by about 10%) and latent heat flux increased (by about 14%) resulting in a balanced budget. Previous flux studies have required similar adjustments, often to ensure that the heat transport in the ocean agreed with oceanographic estimates. The satellite estimates of net SW radiation do not support reduction of the shortwave flux (although they in turn may be over-estimated). However by adjusting the other fluxes in various ways a variety of balanced budgets can be produced. As a result, the various estimates span the range of the values from NWP models and reanalyses[4]. The overall conclusion is that for the global balance, and also for zonal averages, the differences between the various flux products are within the uncertainty of the observational estimates (with only one or two outlying exceptions).

Interestingly, almost all adjustments suggest that the latent heat flux must be increased, toward the higher values for latent heat flux predicted by most models. Such an increase would generally worsen the agreement between climatologies (such as UWM or SOC) and the high quality buoy data; however the latter also depend on transfer coefficients to determine the fluxes. Thus while comparisons of fluxes from ship observations and buoys suggest that

---

3 Here we will refer to *in situ* climatologies produced at the University of Wisconsin-Milwaukee (UWM), the Southampton Oceanography Centre (SOC) and the Institut für Meereskunde, Kiel (IfM).

4 The reanalyses considered were the ECMWF 15 year reanalysis (ERA15), the NCEP/NCAR reanalyses (NCEP1 and NCEP2) and the Goddard GEOS reanalysis (GEOS1).

NCEP1, NCEP2 and ERA15 evaporation may be too large, we can not be certain. This also reflects the uncertainty in our knowledge of the strength of the hydrological cycle. New satellite based products may help through better estimates of precipitation (although there is a lack of verification data) and global estimates of evaporation. With regard to the latter, the HOAPS climatology[5] suggests that *in situ* climatologies under-estimate evaporation in the Southern Ocean. Further development of these satellite products is needed, for example, the HOAPS latent heat flux is biased low in convective tropical regions.

Other significant differences exist between the different products on a regional level and in some cases the cause can be identified. For example, the reanalyses all appear to have substantial errors in surface radiation fields, reflecting problems both with regard to clouds and clear sky radiation. NCEP1 had errors in the downward shortwave and the albedo which tended to cancel. A major shortcoming for ERA15, GEOS1 and NCEP2 appears to be a lack of low-level stratiform clouds over the eastern subtropical oceans. Satellite Surface Radiation Budget (SRB) short wave radiation appears to have more realistic patterns than the reanalyses, but may have a positive bias. Even the estimates from crude *in situ* parameterisations are in better agreement with SRB than are the models.

Choice of Product

Considering the recent, global climatologies based on *in situ* data, UWM and SOC are similar. The SOC climatology gives marginally better agreement with buoy measurements and may be preferred for regional studies. However for a balanced global net heat flux (e.g. for forcing models) then the tuned UWM fluxes must be preferred; these also cover a longer time period. More use should be made of the UWM sensitivity fields for producing new versions of the constrained fluxes. In all cases the sampling density in any particular region must be considered when using these products.

Covering the Atlantic Ocean only, the IFM climatology has used a better mapping technique, has good error estimates, and implies a realistic meridional heat transport. The latter is due to the choice of radiation parameterisations and transfer coefficients which are not generally used for open ocean regions (but are within the bounds of uncertainty). Further verification of this new product is desirable.

The NCEP1, NCEP2 and ERA15 reanalyses appear to have more realistic fluxes than the earlier GEOS1 reanalysis (the GEOS data assimilation system has since been substantially developed). At present the NCEP1 reanalysis covers the longest period and is the most widely distributed and studied. It has the most realistic oceanic low-level stratiform clouds. NCEP2 corrected mistakes in NCEP1 and used improved short wave radiation and boundary layer parameterisations. Its precipitation patterns, downward short wave radiation at the surface, and equatorial wind stress appear superior to NCEP1, but its sensible heat flux over the oceans is lower than all other estimates. ERA15 is the highest resolution reanalysis now available. It appears to produce the most realistic precipitation patterns of the four reanalyses and has more realistic surface downward short wave radiation than NCEP1. Concerning month-to-month variability in surface fluxes, GEOS1 had, in general, the smallest variability, NCEP2 tends to have the largest.

For most of the world ocean, satellite based products are capable of providing much better time and space resolution, and more consistent accuracy, than fluxes based on *in situ* observations. With passing years the restricted time period for which data is available is becoming less of a disadvantage. The surface shortwave flux fields are superior to the model estimates; and presumably also to *in situ* estimates. However on a regional basis the difference between satellite and ship estimates is similar in magnitude to the difference between different satellite fields. The satellite longwave estimates remain less reliable. Satellite scatterometer wind data has significantly less scatter than ship observations. Satellite estimates of the latent

---

[5] Hamburg Ocean Atmosphere Parameters and fluxes from Satellite data.

heat flux have improved substantially but problems still remain in some regions. Satellite precipitation estimates have not been evaluated here other than to review the various comparison projects that have been undertaken by others; lack of verification data is a major problem. For the future, satellite products should exploit combinations of different sensors, preferably flown on a single platform (as for example, the TRMM mission). Combined satellite and *in situ* data products have the potential to provide a data set which is independent from the models.

In summary, all the various flux products that were reviewed in this report have different advantages and disadvantages. However, it is also evident that the various flux products are complementary rather than being in competition, and that the ensemble of air-sea flux estimates that they represent cover the time and space scales that ocean science studies are investigating.

Specific Recommendations

The main conclusions from the Working Group's endeavours appear throughout the body of the report, are summarised in the final Chapter, and have been described above. At its final meeting, the Working Group also adopted the following specific recommendations:

❍ Reanalyses should be performed every 5-10 years by more than one centre; adequate resources should be provided to the reanalysis efforts to improve their surface fluxes, to carry out and evaluate the reanalyses and to ensure that they are easily available to the entire scientific community. Surface fields should be output every 3 hours .

❍ Evaluation of the surface fields and fluxes from global operational NWP systems will benefit future reanalyses as well as provide critical guidance and product uncertainty estimates to users of these flux products. The WGNE's plans to archive and evaluate the surface fields and fluxes from a number of global NWP systems should therefore be supported.

❍ A network of high quality “flux reference platforms” (combination of long-term moorings and ships) should be established to deliver highly accurate values of stress and all components of the air-sea heat fluxes for, *inter alia*, verification of surface fields and flux estimates from satellites and models, and the long term calibration of satellite sensors.

❍ There is continuing need to compare and assess the quality of fluxes derived from various sources, and to evaluate the parameterisations used. Encouragement should be given to efforts to enhance the reliability of momentum, net heat and freshwater fluxes by combining the best estimates from these various sources.

❍ Support should be provided for the continuing assembly of flux and flux-related data sets (in particular Voluntary Observing Ship-based collections such as COADS and other historical data). Continued efforts are needed to remove non-stationary observational biases in historical data. Basic meteorological variables should be included as well as uncertainties, error estimates and adequate documentation for all flux data sets. A catalogue of flux data sets should be maintained on the Internet.

❍ Research and field experiments are needed to improve boundary layer parameterisations and bulk formulae, especially in regions where our physical understanding is poor. Adequate resources for complete analysis of the resulting data are necessary to realise the full benefits of the field experiments.